Spring 2014

Characterization and Tool Development for Studying Transcriptional Activation by NFATc2 and cJun

David McSwiggen
University of Colorado Boulder

Characterization and Tool Development for Studying Transcriptional Activation by NFATc2 and cJun

David McSwiggen

An Honors Thesis submitted to the
University of Colorado
Department of Chemistry and Biochemistry
March, 2014

Defended April 3, 2014

Thesis Advisor: James Goodrich | Biochemistry

Defense Committee: Joseph Falke | Biochemistry
James Goodrich | Biochemistry
Greg Odorizzi | MCDB

Abstract

The underlying processes by which gene expression is regulated is crucial to understanding the onset and progression of many diseases, most notably cancer. RNA Polymerase II transcribes the mRNAs destined to make proteins, but there is an abundance of protein factors which regulate, recruit, inhibit and control Pol II to achieve tight control over which genes are expressed and when. This thesis focuses on two families of transcriptional activators: the NFAT family, and the AP-1 family. Both families of transcriptional activators have been found to have a role in the onset and growth of cancers, but the mechanisms remain unclear.

First, the role of NFATc2 in the cellular response to environmental carcinogens is examined. It is found that cadmium exposure induces calcium oscillations in a breast cancer cell line. This, combined with data from gene reporters and endogenous gene expression, leads to a hypothesis that cadmium may activate NFATc2-controlled transcriptional programs.

Second, the development of two tools which probe different aspects of transcriptional regulation is presented. The first tool, NFATc2-APEX2, uses specially restricted promiscuous labeling to detect proteins which associate with NFATc2. If successful, APEX2-labeling has the potential to detect previously unidentifiable protein interactions.

Finally, this thesis discusses the characterization of a DNA aptamer selected to bind cJun/cJun homodimers selectively and with high affinity, a second tool for understanding transcriptional regulation. The binding affinity of the aptamer for the cJun/cJun homodimer has been determined, along with the aptamer secondary structure, the aptamer interaction regions, and the minimal domain required for high-affinity binding.

Acknowledgments

I would first like to thank Jim Goodrich and Jen Kugel for all of their support their guidance, and their trust. It is a rare thing, indeed, to find people who are so open to letting their students run with a new idea (however bad it might be). I have learned a lot; as much from my failures as my successes. I thank them for fostering an environment that supports this type of exploration and learning.

There are lots of people who have helped me with this research over the past few years. Thanks especially to Ryan Walters, Kurt Cox, and Tereza Ormsby, for their mentorship and patience as we waded through experiments together. I cannot imagine better teachers. Thanks to the rest of the members of the Goodrich Lab, past and present: Petro Yakovchuk, Steve Ponican, Becca Blair, Rob Abrisch, Abby Horn, Joe Cardiello, Vijay Sekaran, and Yogitha Pazhani, and Tess Eidem. I appreciate all of their guidance, and the science we've shared together.

Thanks to all the members of other labs who have helped me with various aspects of my projects. To Genevieve Park, Kevin Dean, Ali Young, and Kyle Carter for their help on all things microscope related. To Jeff Martell and Monica Neugebauer for acquisition and troubleshooting of APEX. To Zach Poss, Ben Allen, and Jack Lin for their advice.

Thanks to Jim and Jen, Ryan, and Jim McSwiggen for their critiques of this manuscript.

A special thanks to my family, especially to my wife who has put up with many late nights and missed dinners in the name of science.

Finally, thanks to the Biological Sciences Initiative, UROP, and Michael and Peggy Touff, whose financial contributions made research possible for me.

Table of Contents

List of Figures

Figure

Chapter 1: An Introduction to Transcription and Transcriptional Regulation in Eukaryotes

Transcriptional regulation can be thought of as a problem in information transfer: There are approximately 20,000 protein-coding genes contained within the human genome, but only a small fraction of these are expressed in any given cell[1]. Contained within this genome are the instructions to make neurons, bone cells, skin cells and any other type of tissue a human needs. But neurons don't need to be able to produce the hydroxyapatite required for bone production, and osteoblasts have no need for the proteins that produce gastric juices. As such, intermediary steps that allow a cell to produce one set of genes over another are vitally important to a successful organism. An understanding of how the information encoded in the genome is expressed in a regulated manner is crucial in designing and improving current treatments for a wide variety of diseases.

RNA Polymerase II and the general transcription factors

Transcription, the process by which RNA is synthesized from a DNA template, is an absolutely essential biological process which is conserved in all known domains of life[2]. This process occurs through the successive incorporation of nucleotide triphosphates (NTPs) catalyzed by the active site of an RNA polymerase, and is energetically driven by the release and rapid degradation of pyrophosphate as each nucleotide is incorporated into the nascent RNA strand[3]. A high degree of message fidelity is achieved through hydrogen bonding interactions between the template strand of the DNA and the incoming NTP, coupled with efficient proofreading[4,5].

RNA Polymerase II (Pol II) is one of three RNA polymerases in eukaryotic organisms. Its function is to transcribe messenger RNA (mRNA) from protein coding sequences in the genome, as well as to transcribe certain snRNAs, mircoRNAs, and newly identified classes of long non-coding RNAs[1]. Pol II consists of twelve distinct protein subunits (RPB1 through RPB12). Together they comprise distinct structural regions that include: (1) the active site for DNA-templated RNA synthesis, (2) a clamp to keep the DNA in the active site, (3) a funnel through which incoming NTPs can enter the enzyme active site, and (4) a "rudder" which splits apart the growing RNA strand from the DNA template[1,6]. The carboxyl terminal domain of RPB1 is composed of up to 52 heptad repeats (YSPTSPS), which are heavily post-translationally

modified, and which are required for the successful production and processing of a mature mRNA[7, 8].

Although Pol II itself can catalyze the synthesis of RNA transcripts, early experiments showed that it lacks specificity for promoter regions, and cannot itself initiate RNA synthesis at a precise location. To do so, it requires a set of nuclear proteins now known as the general transcription factors (GTFs; TFIIA, TFIIB, TFIID, TFIIE, TFIIF, TFIIH). These factors contain a host of activities such as promoter-specific binding and helicase activity which assist in the formation of a stable pre-initiation complex (PIC), DNA melting, and productive transcription initiation[2,9].

PIC assembly begins with the association of TFIID, a massive multi-protein complex consisting of TATA-binding protein (TBP) as well as the TBP Associated Factors (TAFs) (Figure 1.1)[5,10]. Assisted by TFIIA, TFIID binds to and bends the DNA via the TBP subunit, which recognizes the TATA-box, an element of the core promoter ~30 nucleotides upstream of the transcription start site[11]. The TAF subunits of TFIID also associate with promoter elements downstream of the TATA box. It is thought that the bent DNA facilitates the association of Pol II and other GTFs including TFIIB, TFIIE, TFIIF, and TFIIH[12]. Upon binding, the helicase activity of TFIIH hydrolyzes ATP in order to melt a bubble in the DNA, exposing the template strand[13,14]. With the template exposed, Pol II can initiate transcription, but often multiple short abortive products are produced before the polymerase can escape the promoter to transition from an initiation complex into an elongation complex. Once promoter escape occurs, most of the GTFs will dissociate from Pol II and the promoter region, with the exception of TFIIF–which remains associated with Pol II throughout elongation–and TFIID, which remains at the promoter and undergoes another round of PIC assembly[12,15].

Promoter architecture and the regulation of transcription in mammals

The architecture of a eukaryotic promoter consists of a number of cis-acting elements which fall into the broad categories of either the core promoter, proximal regulatory elements, or enhancers[15]. The core promoter contains the TATA-box, the transcription start site, and a downstream promoter element, allowing for the association of Pol II and the GTFs[12]. The proximal regulatory elements flank the core promoter for a few hundred base pairs in either

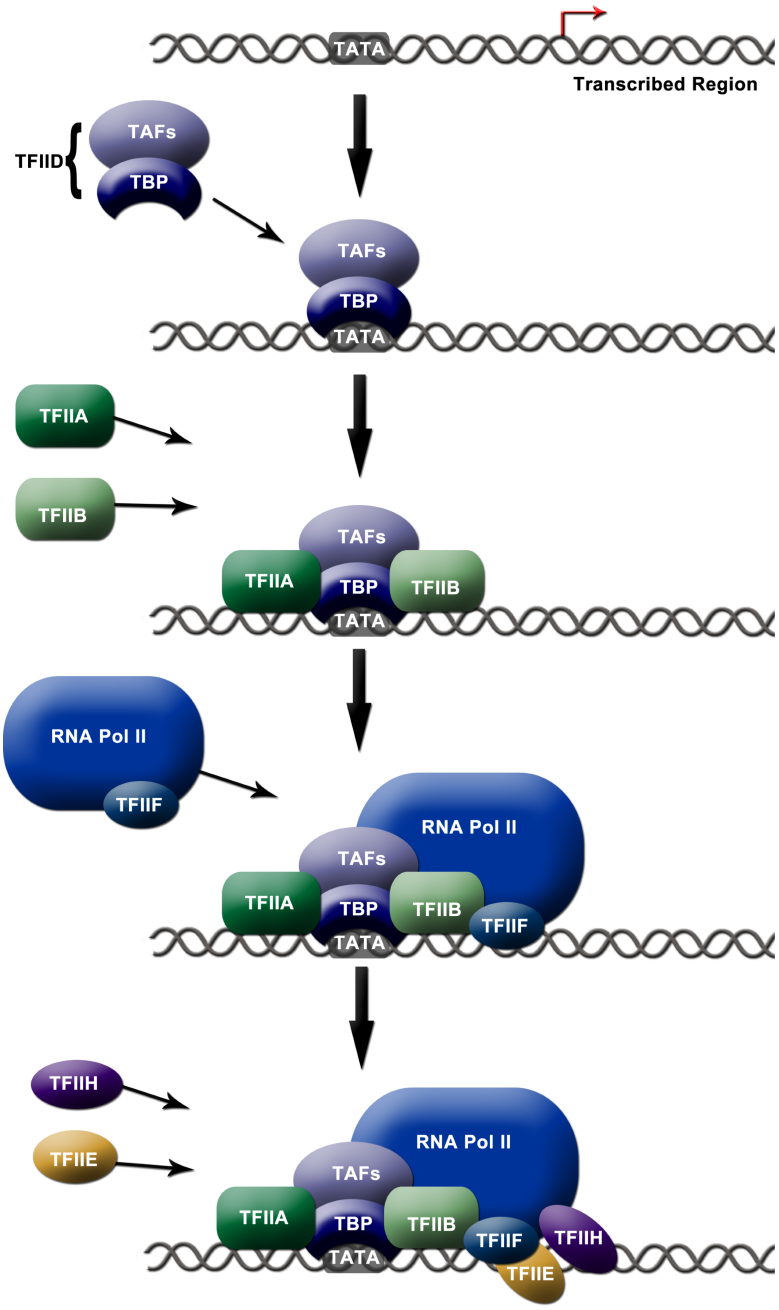


Figure 1.1: A general schematic for the assembly of the PIC. The TATA box, 34 base pairs upstream of the transcription start site (red arrow) is bound by TBP, which is in turn bound to the TAFs and other factors that comprise TFIID (not shown). TFIIA associates and stabilizes the complex. Pol II and TFIIF are recruited to the transcription start site, followed by TFIIE and TFIIF, which melts the DNA to expose the template strand.

direction, and contain recognition elements to which transcription factors (e.g. activators and repressors) may bind. Proximal regulatory elements are often characterized by a high degree of dependance on distance from the core promoter. Enhancers, and other distant regulatory elements, also allow for the binding of transcription factors, but in contrast with the proximal regulatory elements, they are often thousands or tens of thousands of bases away from the core promoter. Enhancers are thought to function by looping out the intervening DNA to allow enhancer-bound transcription factors to interact with factors bound to proximal regulatory elements and the core promoter. As such these enhancer sequences can be functional even when moved to different locations on the chromosome[16].

Transcription factors are proteins that bind to specific DNA sequences in order to regulate the expression of a given gene[17]. This generic term encompasses any factor which binds DNA to regulate gene expression, despite many different functions and mechanisms of action. Some transcription factors aid in the recruitment and assembly of the general transcription machinery, whereas others may interact with the PIC to change expression levels. Often, transcription factors bound to enhancer elements alter transcription levels through interactions with bridging protein cofactors that don't bind DNA directly; for example, the massive protein complex Mediator (Figure 1.2)[18, 19].

Transcriptional activators are combinatorial in nature: multiple activators work together in highly specific combinations to control transcription of individual genes. For example, transcriptional activators A and B may regulate a distinct set of genes compared to transcriptional activators A and C[20]. This combinatorial nature extends to far more than two transcription factors at any particular gene, and many factors are context and cell-type dependent, contributing to the complexity of understanding transcriptional regulation[21].

Transcription is also regulated by modulating the accessibility of promoter DNA to Pol II and the GTFs[22]. The higher order compaction of DNA begins by winding the DNA around histone octamers comprised of two copies each of H2A, H2B, H3 and H4[23]. Regions of DNA in the transcriptionally active euchromatin state contain histones with specific post-translational modifications (e.g. methylation and acetylation) on their tails that facilitate recruitment of Pol II and the GTFs to the core promoter[24]. Regions of the DNA in the heterochromatin state are

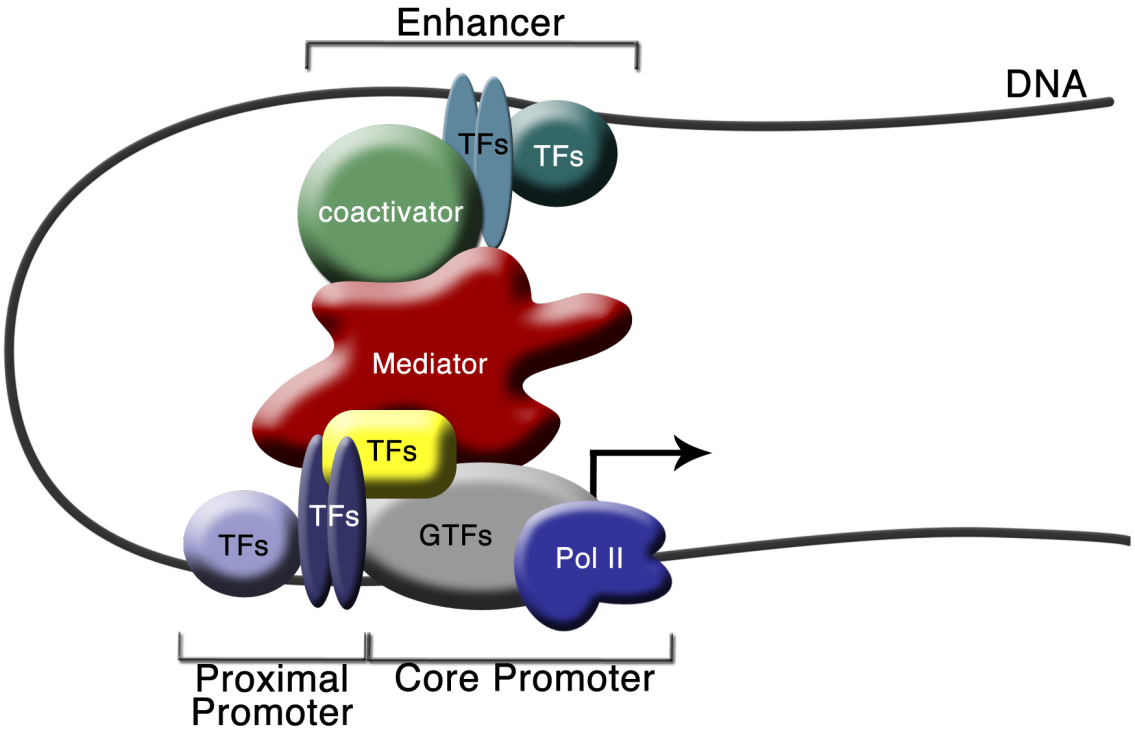


Figure 1.2: The architecture of a mammalian promoter. RNA Polymerase II and the GTFs bind elements that make up the core promoter, present at all genes. The proximal promoter elements can be found within a few hundred bases of the core promoter, and contain recognition sequences for transcription factor binding. Enhancer elements can be found thousands to hundreds of thousands of bases away from the transcription start site, and often tune transcription levels by interacting with protein coactivators and/or mediator.

characterized by repressive histone modifications, inhibiting transcription by preventing Pol II and other proteins access to the necessary elements to initiate transcription. Post translational modifications to the amino-terminal region of histones allow the association of specific protein factors to facilitate these regulatory processes[25, 26].

The AP-1 Family of transcriptional activators

The Activator Protein 1 (AP-1) family of proteins is a large family characterized by their basic leucine zipper (bZip) domain, so named for the five leucine repeats evenly spaced along an α-helix, which allow for dimerization through hydrophobic and electrostatic interactions[27] (Figure 1.3). The act of dimerization in AP-1 proteins is cooperative with DNA binding, such that the binding of DNA induces the formation of the coiled coil of α-helices (Figure 1.3), and greatly increases the binding affinity of each subunit for the other in the dimer pair. All AP-1 dimers contain a DNA binding domain that recognizes the same consensus TPA responsive element (TRE, consensus 5'-TGA(C/G)TCA-3'), but the specific components of the dimer pair alters the activity and specificity of the dimer for regulating promoter-specific transcription[28]. A challenge shared when studying AP-1 proteins, or other transcriptional activators, is the dynamic nature of their activation domains that makes them highly resistant to crystallization or structure determination by NMR. This makes even more difficult the understanding of exactly how these different dimer combinations achieve this specificity.

The AP-1 family can be subdivided into members of the Jun, Fos, ATF and MAF subfamilies, each with specific characteristics. Jun proteins have the ability to form homodimers and to dimerize with other Jun family members, in addition to their ability to form heterodimers with Fos family proteins. Fos proteins, on the other hand, cannot homodimerize, and instead can only heterodimerize complexes with Jun, ATF, or MAF proteins[29]. AP-1 proteins are induced by a variety of extra- and intracellular factors such as the tumor inducer 12-O-tetradecanoylphorbol-13-acetate (TPA) as well as external growth factors and cytokines which activate the mitogen-activated protein kinase (MAPK) signaling cascade. These signals are integrated primarily by Jun N-terminal kinase (JNK), p38, and extracellular signal regulated kinase (ERK). They act by phosphorylating AP-1 proteins, thereby altering the proteins' stability, specificity, and activity[30, 31].

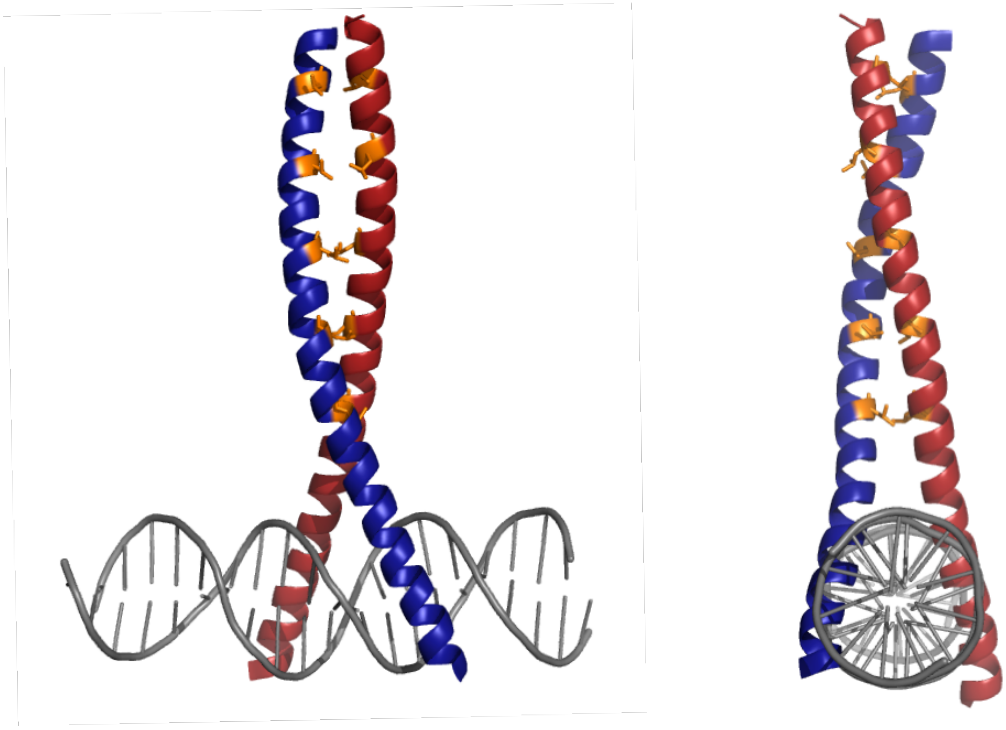


Figure 1.3: The structure of AP-1 proteins. A crystal structure of the basic leucine zipper region of a cFos (red) and cJun (blue) dimer. The leucine repeats (orange) can be clearly seen interacting with each other as the DNA binding domain binds the canonical sequence 5'-TGA(G/C)TCA-3'. Glover, J. N., and Harrison, S. C. Crystal structure of the heterodimeric bZIP transcription factor cFos-cJun bound to DNA, *Nature* **373**, 257-261 (1995).

AP-1 proteins help to govern a vast array of different cellular processes, many involved with cell cycle regulation, growth, and differentiation. The family has long been known to have a role in tumorigenesis, but more recently it has been shown that certain AP-1 family members may act as tumor suppressors[29]. Understanding the exact role of any one dimer pair remains a limiting factor in the ability to target AP-1 proteins for therapeutic research.

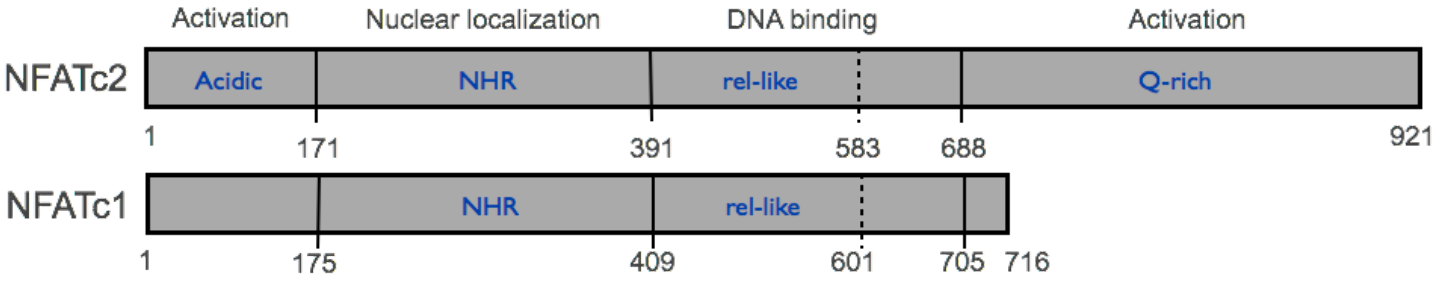
The NFAT family of transcriptional activators

Proteins in the nuclear factor of activated T cells (NFAT) family were first characterized for their role in T cell differentiation during an immune response[32]. Under normal conditions, NFAT proteins remain excluded from the nucleus and in a hyper-phosphorylated state. Efflux of calcium from the endoplasmic reticulum causes activation of calmodulin (a calcium binding protein) and calcineurin (a phosphatase). Activated calcineurin dephosphoyrlates key residues on NFAT, which exposes a nuclear localization signal and allows NFAT translocation into the nucleus[33]. This process is reversible by active kinases within the nucleus, such as casein kinase I or dual specificity tyrosine phosphorylation regulated kinase 2. These kinases work to re-phosphorylate NFAT proteins, causing export of NFAT back out into the cytosol[34].

There are five members of the NFAT family, each of which has multiple alternative splicing variants. The best characterized of these are the proteins NFATc1 and NFATc2. Though structurally distinct, NFATc1 and NFATc2 contain two conserved regions common to their function[35] (Figure 1.4A). The NFAT homology region (NHR) contains phosphorylation sensitive docking sites for calcineurin, as well as the nuclear localization signal for NFAT trafficking. The Rel-like domain contributes to the specificity for DNA binding[34,35].

NFAT proteins also have been found to play a role in cancer development. Biopsies of invasive human carcinomas show a significant increase in NFATc1 and NFATc2 expression[36]. These transcriptional activators drive expression of factors such as Cox2[37] and autotaxin[34], both of which increase cell motility and drive cell invasion. Even still, the mechanisms by which NFAT interacts with other transcription factors in promoting cancer proliferation and invasion remains unclear.

A)



B)

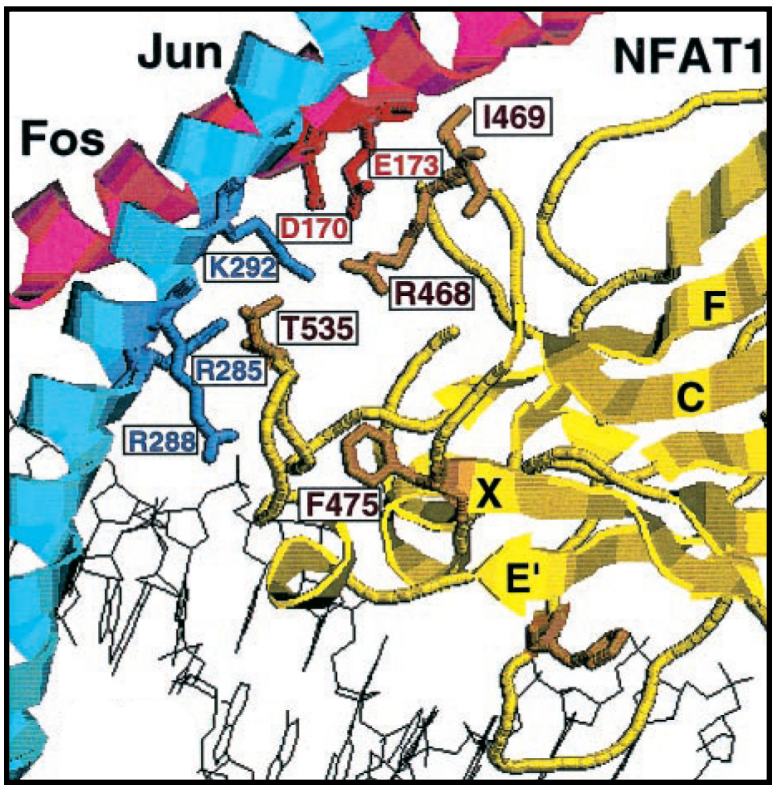


Figure 1.4: The structure of NFAT proteins. **A)** NFATc2 and NFATc1 have two homologous domains: the NFAT Homology region (NHR) that contains the docking site for calcineurin, and the Rel-like domain involved in DNA binding. **B)** A crystal structure of the DNA binding domains of AP-1 and NFAT1 (NFATc2), showing how the proteins can form a stable ternary complex when bound to DNA. Macián, F., López-Rodríguez, C. & Rao, a. Partners in transcription: NFAT and AP-1. *Oncogene* **20**, 2476–89 (2001).

NFAT and AP-1 proteins cooperate to activate transcription

No transcription factor works alone to drive gene expression, but instead works in concert with hosts of other transcription factors, producing the complex control over cell fate that we observe. During T cell development, the co-stimulation of NFAT and AP-1 proteins leads to cooperative binding at the promoter of the IL-2 (interleukin-2) gene, which encodes an essential cytokine for the mammalian immune response[38]. NFATc2 and AP-1 proteins form stable ternary complexes with promoter DNA through interactions between the bZip domain of the AP-1 proteins and Rel-homology domain of NFAT[39] (Figure 1.4B).

The work of this thesis is motivated by two considerations: First, that both AP-1 family proteins (and in particular cJun) and NFAT proteins have been shown to have roles in promoting growth, invasion, and angiogenesis in cancerous cells. Second, that the interactions between these two proteins function in a highly cooperative manner to drive gene expression in the immune response. Consequently, we are interested in determining whether similar interactions might govern the fate of cells in breast cancer.

The work outlined in this thesis focuses on the investigation of NFAT and AP-1 proteins, and discovering better tools to monitor their activity. Chapter 2 presents an ongoing project whose aim is to determine whether known environmental carcinogens drive transcriptional programs controlled by cJun and NFATc2. Chapters 3 and 4 present the development of two tools: one that may be useful in determining what other protein factors work with NFAT to regulate gene expression, and one that can distinguish cJun/cJun homodimers from cJun/cFos heterodimers. It is hoped that this work will result in a better understanding of the mechanisms by which these transcription factors contribute to cancer development and growth.

Chapter 2: Investigating cellular calcium levels, NFAT translocation, and transcript levels in response to environmental carcinogens

Introduction

As the incidence of cancers continue to increase—particularly in the Western world—an understanding of the interplay between environmental factors and human health has become more critical to the understanding, diagnosis and treatment of the disease[40]. Cancer is an umbrella term for the atypical division of cells leading to the formation of tumors. It is an incredibly heterogenous disease which can afflict nearly all tissue types, and can have numerous different causes including genome instability and catastrophic DNA damage. The result is a cell or tissue that fails to express the normal safeguards preventing cells from overgrowing their environment[41]. Environmental factors can initiate or exacerbate these effects by altering the normal transcriptional program in the cell, potentially contributing to the proliferation or invasiveness of the cancerous tissue[42]. Two well established groups of environmental factors that promote tumor growth or invasion are polycyclic aromatic hydrocarbons (PAHs) and heavy metals[43, 44].

PAHs are a broad class of organic compounds characterized by highly conjugated systems of aromatic rings. They are the byproducts of numerous production and refining processes such as crude oil refinement, coke oven emissions, motor vehicle emissions, and tobacco smoke[45]. They are metabolized by cytochrome enzymes, such as those involved in oxidative phosphorylation. Metabolism of PAHs, in particular benzo[a]pyrene (B[a]P), forms highly reactive epoxides which can interact covalently with DNA, permanently damaging it. In addition, B[a]P can be metabolized to a quinone which can interfere with the Q cycle and other cellular processes[45].

Heavy metals are another broad class of environmental carcinogens, often appearing in sites where mining and mineral enrichment are prevalent. A notable issue with many heavy metal carcinogens is their tendency to accumulate in tissues, and hence their levels increase as one moves up the food chain. Divalent heavy metals such as cadmium, chromium(VI), nickel(II), cobalt(II), and mercury have all been shown to induce carcinomas in animal studies[46]. One hypothesis suggests that these divalent cations mimic free Ca^{2+} to misregulate signaling pathways activated by calcium, but exact mechanisms have not been elucidated[47].

With both PAHs and heavy metals, the sequence of events from carcinogen exposure to the formation of a cancerous tumor is unclear. However, studies have shown that both B[a]P and cadmium increase levels of free calcium in the cytosol[42-44, 48]. Given the mechanism by which the NFAT family of transcriptional activators are activated in a calcium dependent manner, this led us to wonder if exposure to these environmental carcinogens was initiating the transcriptional programs driven by NFAT. Additionally, recent studies have found that exposure to cadmium activates the MAPK signal cascade and has been shown to increase cJun expression in human prostate cells[46, 49]. Together, these led us to wonder whether NFAT was cooperating with cJun to mediate transcriptional activation in this capacity.

Presented here are the results of experiments monitoring calcium levels, NFAT localization, and transcript levels after breast cancer cells were treated with cadmium and B[a]P. We used a genetically encoded calcium sensor to evaluate the previously published data supporting an increase in cytosolic calcium upon treatment with cadmium and B[a]P. We then characterized the activity of a fusion protein that allows us to follow the translocation of NFATc2 from the cytoplasm to the nucleus in real time. Finally, we examine the endogenous expression levels of mRNAs thought to be activated by NFAT to phenotypically characterize the effect of these carcinogens.

Results

Genetically encoded calcium sensors do not show an increase in cytosolic calcium

We wanted to see if exposure to either B[a]P or cadmium would induce an increase in cytosolic calcium. To this end, we used D3cpV–a genetically encoded calcium sensor consisting of calmodulin flanked by cyan fluorescent protein (CFP) on one end, and yellow fluorescent protein (YFP) on the other[50, 51]. Because of their close proximity, excitation of the CFP can cause fluorescence resonance energy transfer (FRET), transferring energy from the CFP to the YFP and resulting in a photon emission from the YFP in a distance-dependent manner. Upon binding calcium, there is a conformational change in the calmodulin domain, bringing the two fluorescent proteins closer together and changing the FRET between the two. These changes in FRET can be expressed as the ratio of the intensity of the YFP FRET emission to CFP emission.

D3cpV was expressed in the mammary gland adenocarcinoma cell line MDA-MB-231. To test the FRET-based calcium sensor, initial experiments treated cells with ionomycin, an ionophore which floods the cytosol with calcium from the ER[52]. When ionomycin was added to cells in a calcium-free buffer in the presence of a chelator such as EGTA, there was a sharp spike in FRET followed by a return to the basal FRET state, corresponding to a release of calcium from the ER followed by sequestration of the free calcium by EGTA (Figure 2.1A). The buffer was changed to a calcium-rich buffer, and the cells were allowed to recover for approximately 3 minutes. When they were then treated with ionomycin alone, we observed a sustained 3-fold increase in the YFP/CFP FRET ratio. This confirms that the D3cpV reporter is successfully responding to changes in intracellular calcium.

To address whether treatment with carcinogens would lead to an increase in free cytosolic calcium, we treated cells with either cadmium or B[a]P in phenol-free growth medium and imaged them at 2 minute intervals over a 12-hour time course. Contrary to expectation, and to what has previously been reported in the literature, neither B[a]P nor cadmium exposure resulted in a sustained increase in free calcium. Instead, FRET ratios remained relatively constant across the time course (Figure 2.1B, C).

Cadmium-treated cells experience calcium oscillations

During the optimization of the imaging conditions an intriguing result was observed. MDA-MB-231 cells treated with cadmium experienced relatively regular calcium oscillations. To explore this further, we transfected cells with D3cpV and imaged them at 15 second intervals after cadmium treatment. Not all cells experienced these regular oscillations, and the magnitude of the changes varied from cell to cell, but the majority of cells showed calcium oscillations with a period of approximately 4.5 minutes. Neither untreated cells, nor cells treated with B[a]P experience these oscillations (data not shown). Furthermore, cadmum-induced oscillations extended for almost 2 hours after treatment (Figure 2.2).

mApple-NFATc2 is functional and transcriptionally active

Intrigued by the possibility that the calcium oscillations we observed could be a means of regulating NFAT signaling, we wanted to design a fluorescent protein fusion that could be used to monitor the translocation of NFATc2 in cells in real time. mApple seemed to be a fluorescent

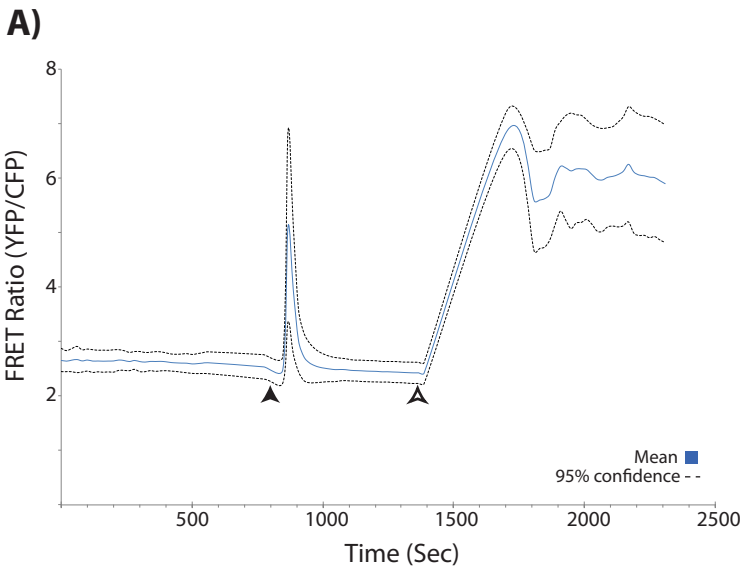
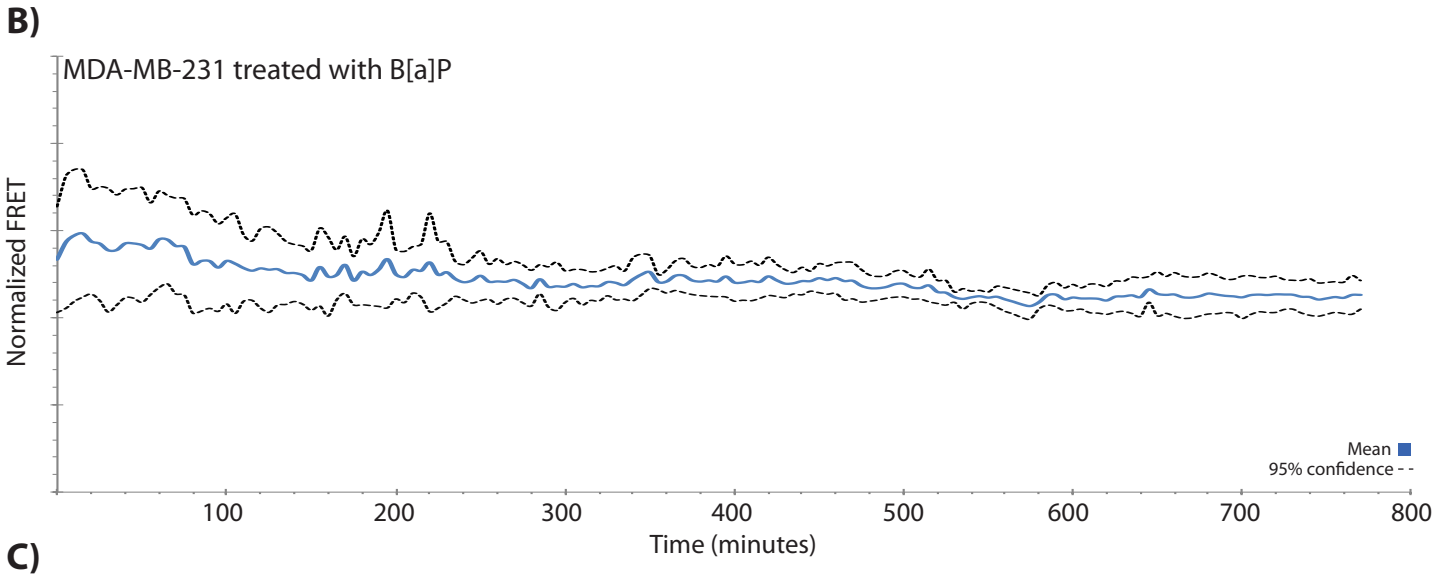
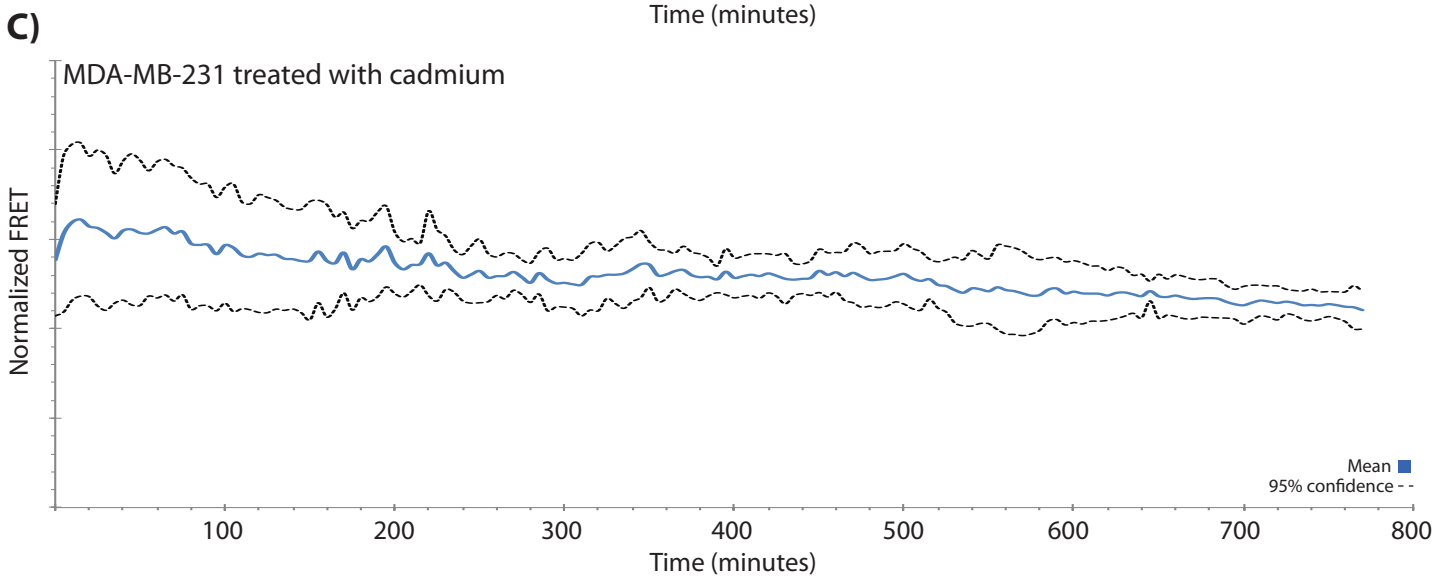


Figure 2.1: FRET sensors detect cytosolic calcium. **A)** MDA-MB-231 cells with a genetically encoded calcium sensor were treated with ionomycin and EGTA in a calcium-free buffer (black arrow), causing release of calcium from the ER. After a buffer exchange, cells treated with ionomycin in a calcium rich buffer experienced an immediate and sustained increase in cytosolic calcium (white arrow).
B and C) MDA-MB-231 cells with a genetically encoded calcium sensor were imaged over night after treatment with B[a]P or cadmium. The mean represents the average of the normalized FRET ratios of 5 cells in one field of view.





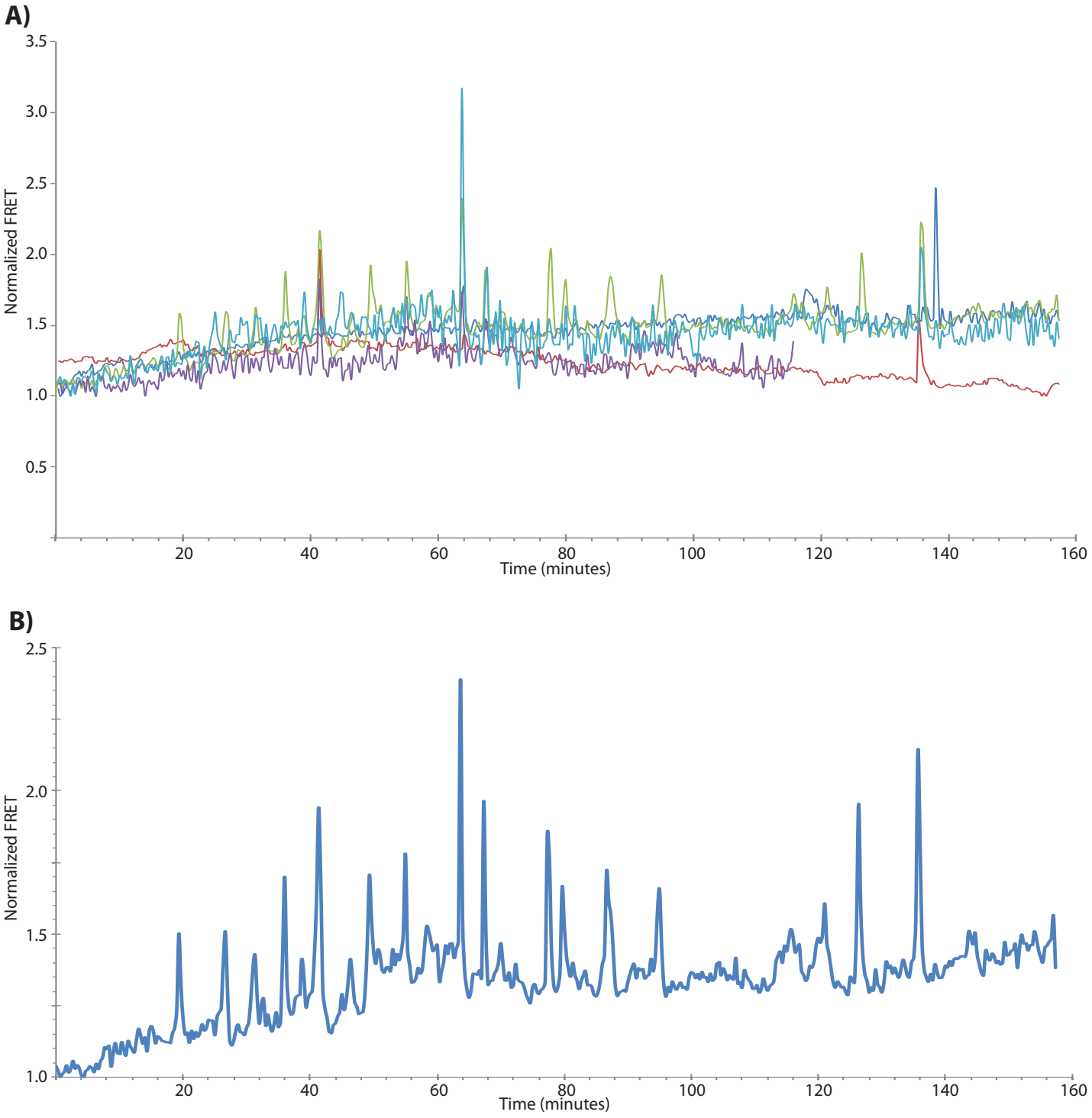


Figure 2.2: MDA-MB-231 cells experience Ca^{2+} oscillations. **A)** The normalized FRET ratios of five cells within one field of view. After treatment with cadmium, three cells experienced oscillations in cytosolic calcium with a period of approximately 4.5 minutes. **B)** The normalized FRET ratios of a single cell after cadmium treatment.

protein ideally suited to the task because of its monomeric character and photostability, as well as an excitation range outside that of CFP and YFP[53]. Fusing mApple to NFATc2 would allow us to examine NFAT translocation as a function of both cytosolic calcium and time, as cells are treated with cadmium or other environmental carcinogens.

mApple was cloned onto the amino terminus of NFATc2, and then assayed for activity to test whether mApple-NFATc2 could still activate transcription at the IL-2 promoter. We chose to use the IL-2 reporter to test activity because NFATc2 is known to act synergistically with cJun to induce high levels of transcription at this promoter. In addition, recent findings suggest that cJun/cJun homodimers bind the carboxyl terminus of NFATc2 in a cofactor interaction in order to get full activation of the gene[54]. Thus, using the IL-2 promoter to test the functionality of mApple-NFATc2 addresses not only DNA binding activity, but also whether fusion perturbs interactions with cJun/cJun homodimers.

We used an IL-2 Luciferase reporter cotransfected into Cos7 cells with either wild-type NFATc2 and cJun, or mApple-NFATc2 and cJun. Transfected cells were stimulated with PMA and ionomycin, then Luciferase activity was monitored (Figure 2.3). The Luciferase data show a three-fold decrease in activity with the mApple-NFATc2 fusion protein compared to the wild-type NFATc2. Despite the decrease in efficacy, however, over-expressing mApple-NFATc2 and cJun together still exhibited synergistic activation compared to either mApple-NFATc2 or cJun alone, suggesting that the reporter is still localizing to the promoter and activating transcription to some degree.

We next asked whether translocation of mApple-NFATc2 into the nucleus might be the limiting factor in IL-2 activation (Figure 2.4A). Cells were stimulated with PMA and ionomycin for 7 hours, just as in the Luciferase assay. Prior to stimulation, cell nuclei were devoid of mApple signal. After stimulation for 7 hours, it appeared that some signal had migrated to the nucleus; however, the amount of nuclear NFATc2 was much lower than expected. The addition of 30mM lithium acetate, which inhibits the kinases that promote NFAT export, showed a much stronger nuclear presence of mApple-NFATc2, although the pattern of fluorescence appeared much more punctate than cells not treated with lithium acetate (Figure 2.4B). These experiments suggest that mApple-NFATc2 does enter the nucleus upon stimulation, but perhaps it is exported

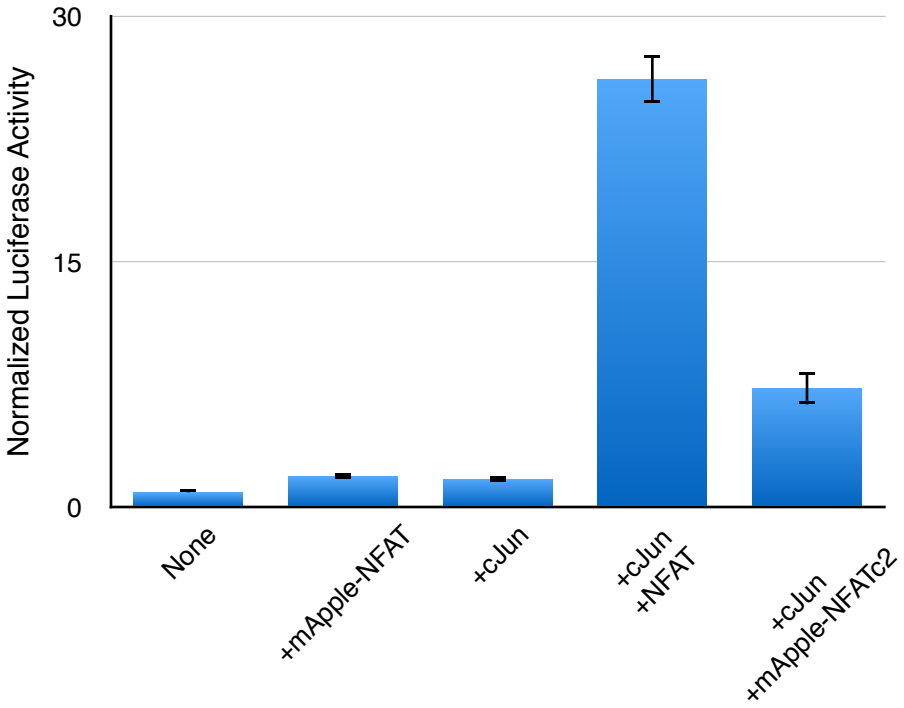


Figure 2.3: mApple-NFATc2 is transcriptionally active. Cos7 cells were cotransfected with plasmids coding for IL-2-Luciferase and Renilla-null. Relative Luciferase activity was determined by dividing the Firefly Luciferase activity by the Renilla Luciferase activity. The data were normalized to the average activity in cells not over expressing either cJun or NFATc2. Data represent the average of three transfections and error bars represent one standard deviation.

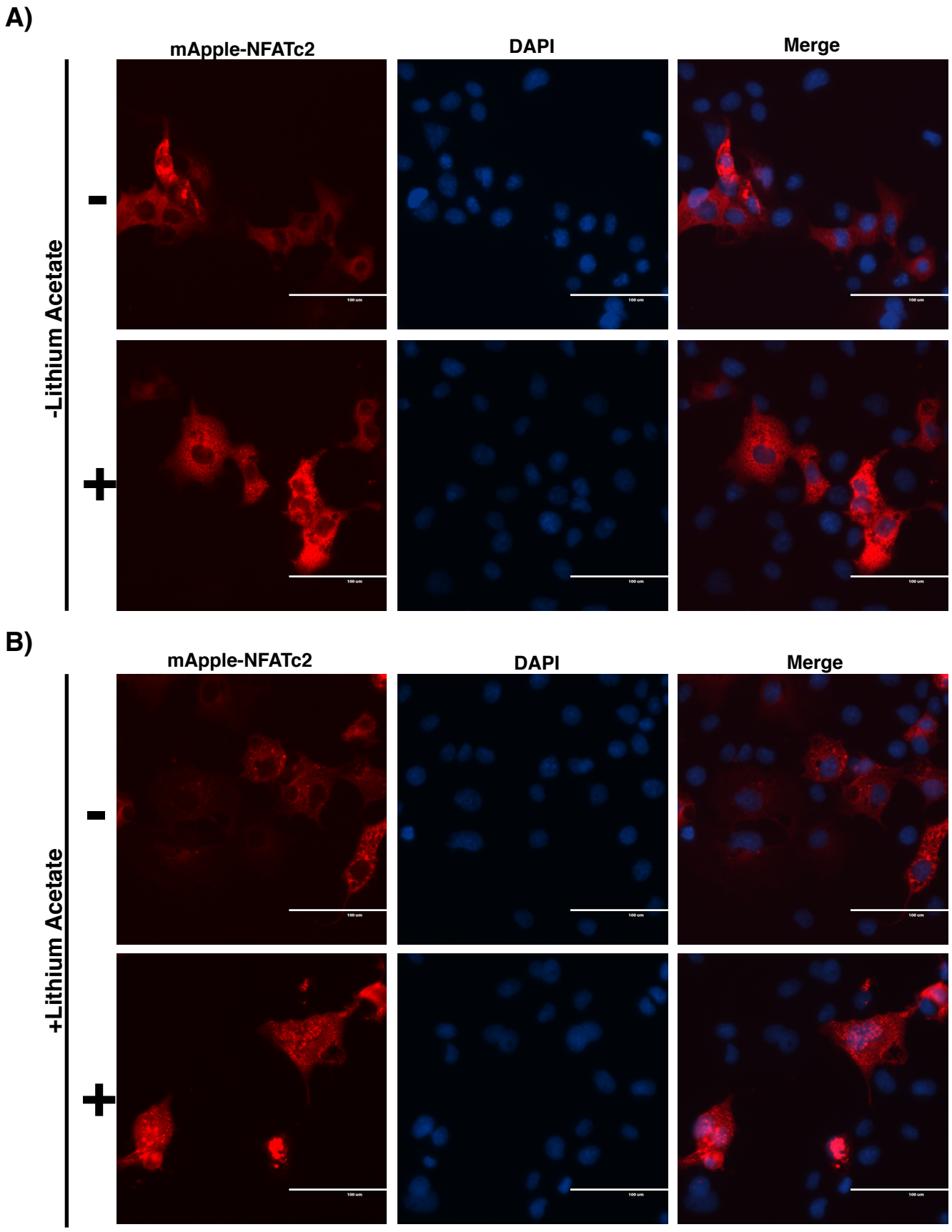


Figure 2.4: mApple-NFATc2 localizes to the nucleus. **A)** Cos7 cells transfected with mApple-NFATc2 were fixed and stained before (-) or after (+) 7 hours of stimulation with PMA and ionomycin. **B)** Cos7 cells transfected with mApple-NFATc2 were treated with Lithium Acetate, then fixed and stained before (-) or after (+) 7 hours of stimulation with PMA and ionomycin. DAPI staining indicates the nucleus, scale bars are 100μm.

more quickly than wild-type NFATc2, leading to a decrease in its transcriptional activity. We are currently testing this construct by live cell imaging in MDA-MB-231 cells to try to evaluate how NFATc2 responds to stimulation.

Cadmium activates Cox2 expression in MDA-MB-231 cells

We also monitored the ability of cadmium to stimulate expression from reporter genes that are NFAT responsive, reasoning that activation of these genes would require NFAT translocation to the nucleus. We chose to look at expression levels of IL-2 and Cox2 using luciferase reporters. IL-2 and Cox2 were chosen specifically because both promoters contain NFAT and AP-1 binding sites that are crucial for gene expression[54, 55].

MDA-MB-231 cells transiently transfected with plasmids encoding either IL-2-Luc or Cox2-Luc were treated with either 3μM B[a]P or 1μM $CdCl_2$ for between 2 and 20 hours. Figure 2.5 shows the normalized Firefly luciferase values. While both IL-2 and Cox2 showed slight increases upon cadmium exposure, the largest increase in expression was only approximately 40% greater than the untreated control. Exposure to B[a]P showed even less effect on the luciferase values with the maximum Cox2 increase only 25%, and no significant IL-2 induction at any time point.

Cadmium activates genes associated with NFATc2 and cJun

We wanted to also assess endogenous mRNA expression to better characterize the transcriptional programs that B[a]P and cadmium are initiating. To accomplish this, we treated MDA-MB-231 cells with 3μM B[a]P or 1μM $CdCl_2$, and for comparison, 1μM ionomycin and 20ng/mL PMA, which robustly activate the NFAT and cJun signaling cascades respectively. After 6 hours, the whole cell RNA was extracted with TRIzol (Invitrogen), and a cDNA library was constructed. Gene-specific primers were then added into a semi-quantitative PCR reaction.

The gene products of Cox2, ENPP2, and TGF-β have all been shown to have a role in tumorigenesis or invasion[41, 56-58], and are driven in part by the activation of cJun, NFATc2, or both. We additionally tested the IL-8 gene as it responds to cJun and NFATc2 stimulation[49], and HSP40 because of its previously established role in the response to cadmium exposure[57].

Duplicates of each condition were not as consistent as we had hoped. However, as we expected, cells stimulated with ionomycin more robustly transcribe Cox2, ENPP2, IL-8, HSP40

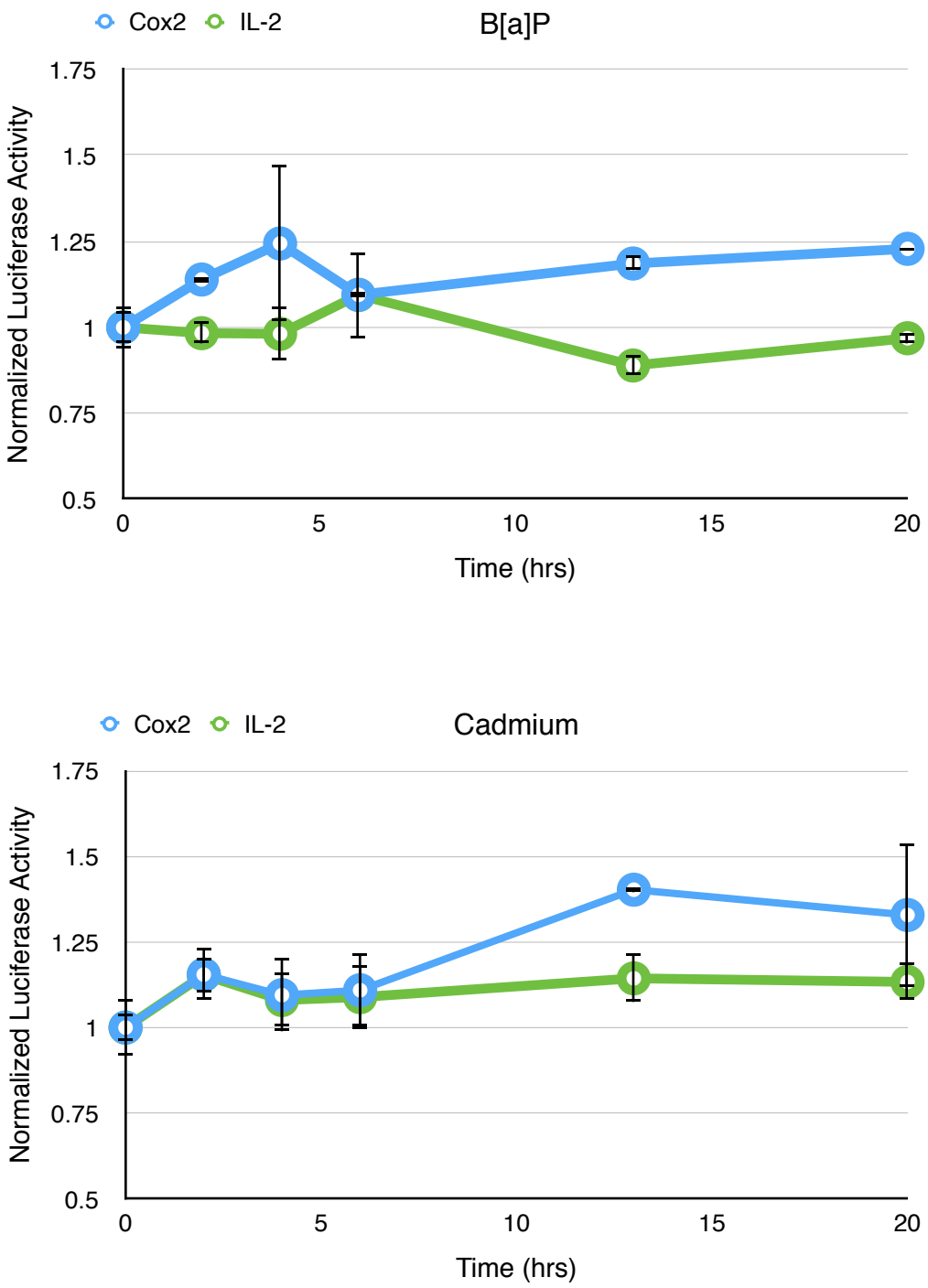


Figure 2.5: Gene reporters in MDA-MB-231 cells. MDA-MB-231 cells were cotransfected with plasmids encoding either Cox2-Luc and Renilla-null, or IL-2-Luc and Renilla-null. Cells were treated with B[a]P or cadmium for 2, 4, 6, 13, or 20 hours. The values represent the luciferase values normalized to the untreated sample. Data represent the average of two samples, and the error bars are the range.

and TGF-β (Figure 2.6). Cells that have been treated with cadmium experience levels of activation similar to inonmycin treatment for Cox2, ENPP2, IL-8 genes, and a stronger induction of HSP40 and TGF-β. In contrast with cadmium, B[a]P does not induce ENPP2, and appears to repress IL-8 expression (Figure 2.6). In either case, the lack of consistency between wells suggests that the experiments should be repeated, and perhaps a more quantitative method such as q-PCR employed.

While these genes only represent a handful of all the potential genes we could possibly examine, they are sufficient to begin to compare the transcriptional effects of B[a]P and cadmium, and whether they activate genes previously published to be responsive to NFAT. Based on these data, it appears that cadmium is the more likely of the two carcinogens to activate cJun or NFAT-mediated transcription. B[a]P clearly induces stress on the cells, as evidenced by the induction of HSP40, but there is no evidence currently that it activates either cJun or NFATc2.

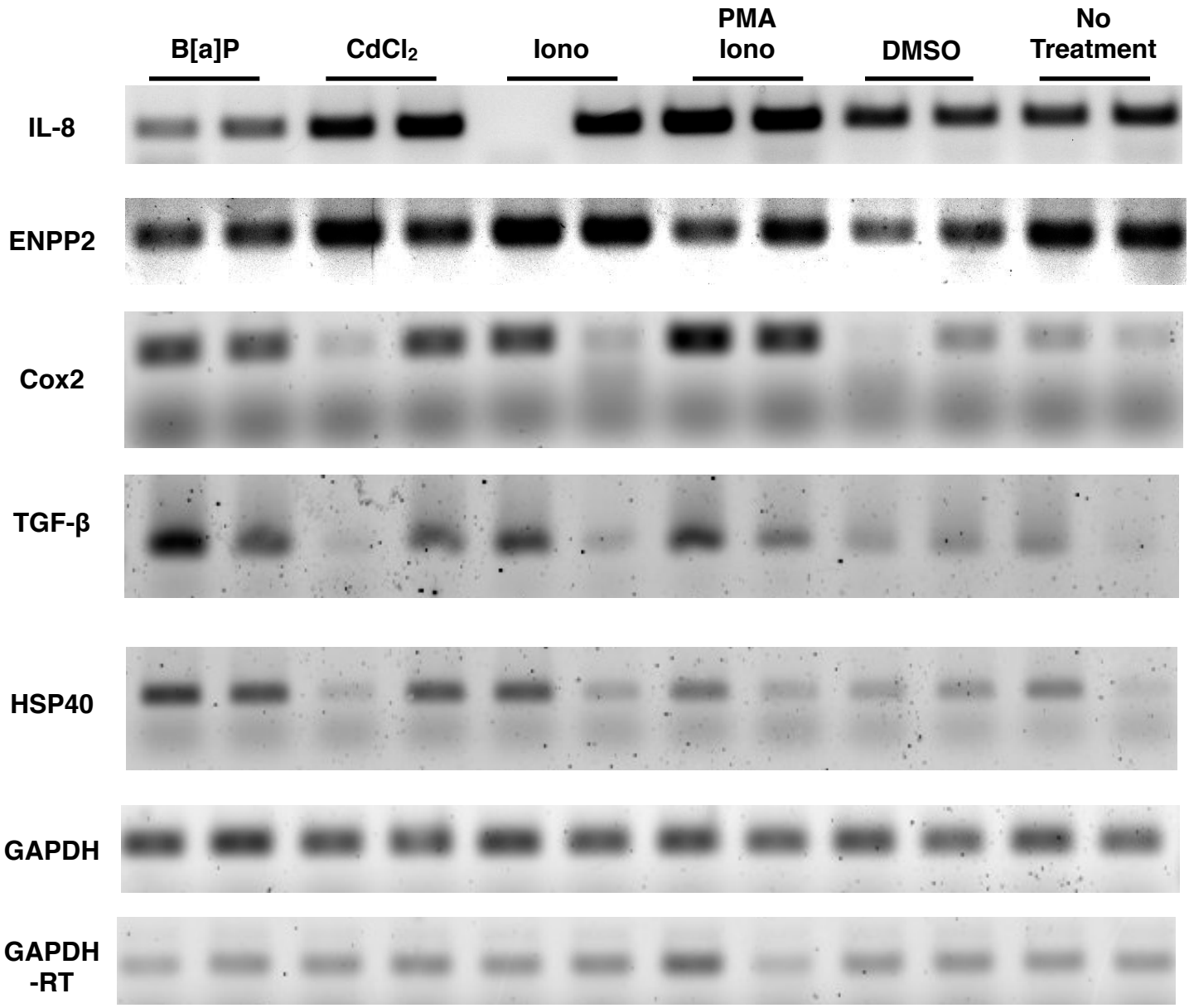


Figure 2.6: MDA-MB-231 cells express NFATc2 and AP-1 inducible genes. Cells were treated in duplicate with B[a]P, $CdCl_2$, ionomycin, or ionomycin and PMA for 6 hours and the whole cell RNA extracted. A cDNA copy was generated using a reverse transcriptase and random hexamers, and then gene products amplified off the cDNA using gene-specific primers. GAPDH -RT samples were treated exactly the same as the other samples with the exception of the exclusion of the reverse transcriptase.

Discussion

Cadmium-induced calcium oscillations may influence NFATc2-controlled genes

Both polycyclic aromatic hydrocarbons and heavy metals such as cadmium have been shown to be carcinogenic compounds, but the mechanisms by which these compounds promote tumorigenesis is still unclear. Given the well established roles that both cJun and NFAT transcriptional activators play in certain cancers, we wondered if either of these compounds would promote either NFAT or cJun activity. In order to probe this question, we first used a genetically encoded calcium sensor to examine how exposure to these compounds affected levels of calcium in the cytosol of a mammary gland adenocarcinoma.

Contrary to previously published results–which used small molecule calcium reporters–we did not see any increase in cytosolic calcium levels in this cell line over similar time periods. In order to address whether these discrepancies are due to experimental methodology, or whether the lack of calcium increase is cell line-specific, one would need to perform both calcium sensing methods in parallel.

We were, however, optimistic about seeing calcium oscillations upon cadmium exposure. The ubiquitous nature of calcium as a second messenger in cells has recently led researchers to question whether the concentration of calcium in the cytosol alone is enough to specifically stimulate a cellular response. Experiments have demonstrated that calcium oscillations contribute to the specificity of signal transduction, such that different frequency oscillations activate different signal cascades[34,59]. For example, T cells exposed to either constant or oscillating calcium concentrations were shown to have a much more robust and selective activation of NFAT under the oscillatory conditions than under conditions of constant calcium. We believe that a similar mechanism of activation may be occurring in these breast cancer epithelial cells, which may help to explain the induction of NFAT responsive genes without the expected steady-state increase in cytosolic calcium.

The induction of genes upon cadmium treatment that are regulated by NFATc2, like Cox2 and TGF-β, provides evidence of that. Thus, changes in the transcriptional program in cadmium-treated cells may be mediated by NFAT activation, and are perhaps also responsive to AP-1 inducible elements. Several of the induced genes have documented roles in carginogenesis,

supporting the model that NFAT, and potentially cJun, activation could be a key player in this process. To fully understand how cadmium works as a carcinogen by changing transcriptional programs, a more thorough examination is needed through ChIP-seq of NFAT, combined with RNA-seq or Pol II ChIP-seq to assess changes in transcription. Only then can we unravel the role that cadmium is playing in NFAT-mediated transcriptional changes. Still, the above results leave us confident that we will find NFAT as a key player in mediating this cellular response.

Materials and Methods

Chemicals and reagents

Cadmium chloride (Sigma Aldrich) was dissolved in ultra-pure water to a stock concentration of 1 M and stored at -20°C. Benzo[a]pyrene (Sigma Aldrich) was dissolved in tissue culture grade DMSO to a stock concentration of 50 mM and stored at -20 °C for up to one month.

Cell culture

Human mammary adenocarcinoma cells (MDA-MB-231) were cultured in Libovitz's L-15 medium supplemented with 10% (v/v) fetal bovine serum, 2 mM L-glutamine, 100 U/mL penicillin and 100 µg/mL streptomycin. Cells were maintained at 37 °C and 0% CO_2 in a humid incubator; they were trypsinized with 0.25% trypsin and subcultured roughly every four days. Genetically-encoded reporters were transfected using TransIT-BrCa (Mirus Bio) or electroporated using the Neon transfection system (Invitrogen) according to the manufacturer's instructions. Cells were treated with $CdCl_2$ by exchanging medium with fresh medium containing 1 µM $CdCl_2$. Similarly, cells were treated with B[a]P or DMSO by exchanging medium with fresh L-15 medium containing 3 µM B[a]P. In all cases, the amount of DMSO in the medium never exceeded 0.3% (v/v). Live cell imaging was done at 37 °C and 80% humidity on a Nikon TE-2000 Widefield microscope (Nikon).

Cercopithecus aethiops kidney cells (Cos7) were cultured in DMEM supplemented with 10% (v/v) fetal bovine serum, 2 mM L-glutamine, 100 U/mL penicillin and 100 µg/mL streptomycin. Cells were trypsinized with 0.25% trypsin, subcultured roughly every two days, and maintained at 37°C and 5% CO_2 in a humid incubator. Genetically encoded reporters were transfected using TransIT-2020 (Mirus Bio) according to the manufacturer's instructions.

Plasmid construction

Plasmid HA-NFATc2 (previously described[60]) was amplified with primers containing XhoI and EcoRI restriction sites, and then digested with those restriction enzymes. Plasmid pcDNA 3.1(-) (Invitrogen) was similarly digested, and treated with calf intestinal phosphatase (CIP). The vector and insert were then combined and ligated. mApple was amplified from pcDNA 3.1(+)-mApple-NLS (A generous gift from Amy Palmer, Boulder, CO) with primers containing a Kozak sequence, and containing XbaI and XhoI restriction sites, then digested accordingly. pcDNA 3.1(-)-HA-NFATc2 was digested with XbaI and XhoI, followed by CIP treatment. The vector and insert were ligated before transformation. All constructs were verified by sequencing.

Luciferase assays

250 ng of each protein expression vector, or blank vector, were transfected into Cos7 cells at 60% confluence in a 12-well plate using TransIT-2020 (Mirus Bio) as per the manufacturer's instructions. 20 hours post-transfection, cells were stimulated with 20 ng/mL PMA and 1 μM ionomycin for 7 hours. After 7 hours, cells were harvested in 200 μL 1x Passive Lysis Buffer (Promega) for 15 minutes at room temperature. Firefly and Renilla luciferase values were determined using the Dual-Luciferase Kit (Promega), and the Firefly values normalized to the Renilla values in each well.

mRNA extraction

MDA-MB-231 cells at 70% confluence were treated for 6 to 12 hours with $CdCl_2$, B[a]P, PMA and Ionomycin, or DMSO. Post treatment, cells were scraped from the plate and spun at 1100x g for 5 minutes. Supernatant was discarded, and the cells were suspended in 200 μL TRIzol (Life Technologies) followed by incubation for 5 minutes. 40 μL of chloroform were added and, after 15 seconds of gentle agitation, the solution was centrifuged at 16,000x g for 15 minutes. The aqueous phase was removed and the RNA was precipitated with 100 μL isoporpanol. The collected RNA was pelleted, and washed with 75% ethanol before suspension in 1x RQ-1 DNase buffer (Promega). Samples were treated with 1 unit of RQ-1 DNase (Promega) for 20 minutes at 37 °C, followed by the addition of 1 μL DNase stop solution and a 10 minute incubation at 65 °C. cDNA was synthesized from the extracted RNA using 1 unit of Multiscribe RT, 10 μM random hexamer primers, and 2 mM dNTPs in 1x RT buffer (Promega).

Samples were incubated at 25 °C for 5 minutes, then at 37 °C for 1 hour before ramping up to 65 °C to inactivate the Multiscribe. Individual mRNA products were detected using custom primers in 30-34 cycles of PCR. PCR products were resolved on a 1.8% agarose gel and detected using fluorescence scanning (Typhoon).

Chapter 3: Capturing transient coactivator interactions at the gene promoter

Introduction

The combinatorial control over gene expression exhibited by transcription factors presents a particularly difficult problem for trying to tease apart how any one gene is regulated. The problem is complicated by the fact that transcription factors can function either by binding elements on the DNA or through protein-protein interactions[16]. At present, methods for investigating these transcription factors in cells are limited to chromatin immunoprecipitation (ChIP) and its derivatives, or co-immunoprecipitation (CoIP) and its derivatives. The former can yield substantial information about the sites at which a protein of interest in bound on the genome, particularly when combined with next generation sequencing technology. Its scope is limited, however, in that the investigator must know the identity of their protein of interest ahead of time. In addition, the ChIP signal decreases if a promoter-associated protein is not directly bound to DNA. CoIP, on the other hand, identifies protein-protein interactions but requires these interactions to be very stable so as to withstand stringent wash conditions. Moreover, CoIPs are usually performed from soluble extracts that exclude factors stably associated with chromatin. We wanted to develop a technique that could help bridge the gap between these two established techniques.

Recently, Dr. Alice Ting and her colleagues reported work on an enzyme that they used for spatially restricted proteomic mapping in mitochondria[61]. They showed that APEX, an enzyme engineered from soybean ascorbate peroxidase (APX), has the ability to utilize hydrogen peroxide to generate aromatic radical species[62]. When given a substrate consisting of biotin conjugated to a phenol group, APEX generates a reactive phenoxyl radical with a half-life of approximately 1 millisecond. The reactivity of the phenoxyl radical is such that it can covalently interact with the exposed aromatic residues of proteins within a tight radius of where the radical is created (Figure 3.1). The promiscuous labeling of nearby proteins with a covalent biotin tag allows researchers to use simple straptavidin enrichment, in tandem with mass spectrometry, to identify the labeled species. The Ting group showed that by genetically targeting APEX to the mitochondria, they could achieve labeling of upwards of 96% of the known mitochondrial proteome[61].

The second generation of APEX, APEX2, was designed to be more versatile by having less of a dependance on heme incorporation to achieve a functional protein. Using this improved enzyme, Dr. Ting and colleagues have shown that APEX2 can be used to identify the proteome of systems that are not enclosed by a membrane (unpublished data). The fast labeling time of the enzyme coupled with the relatively nontoxic substrate seem to make APEX2 an ideal candidate for identifying proteins associated with specific transcription factors that might evade detection through ChIP or CoIP. We hypothesized that a genetically encoded APEX2 fusion to the transcription factor NFATc2 would allow us to identify protein partners of NFATc2 that mediate its function as a transcriptional activator.

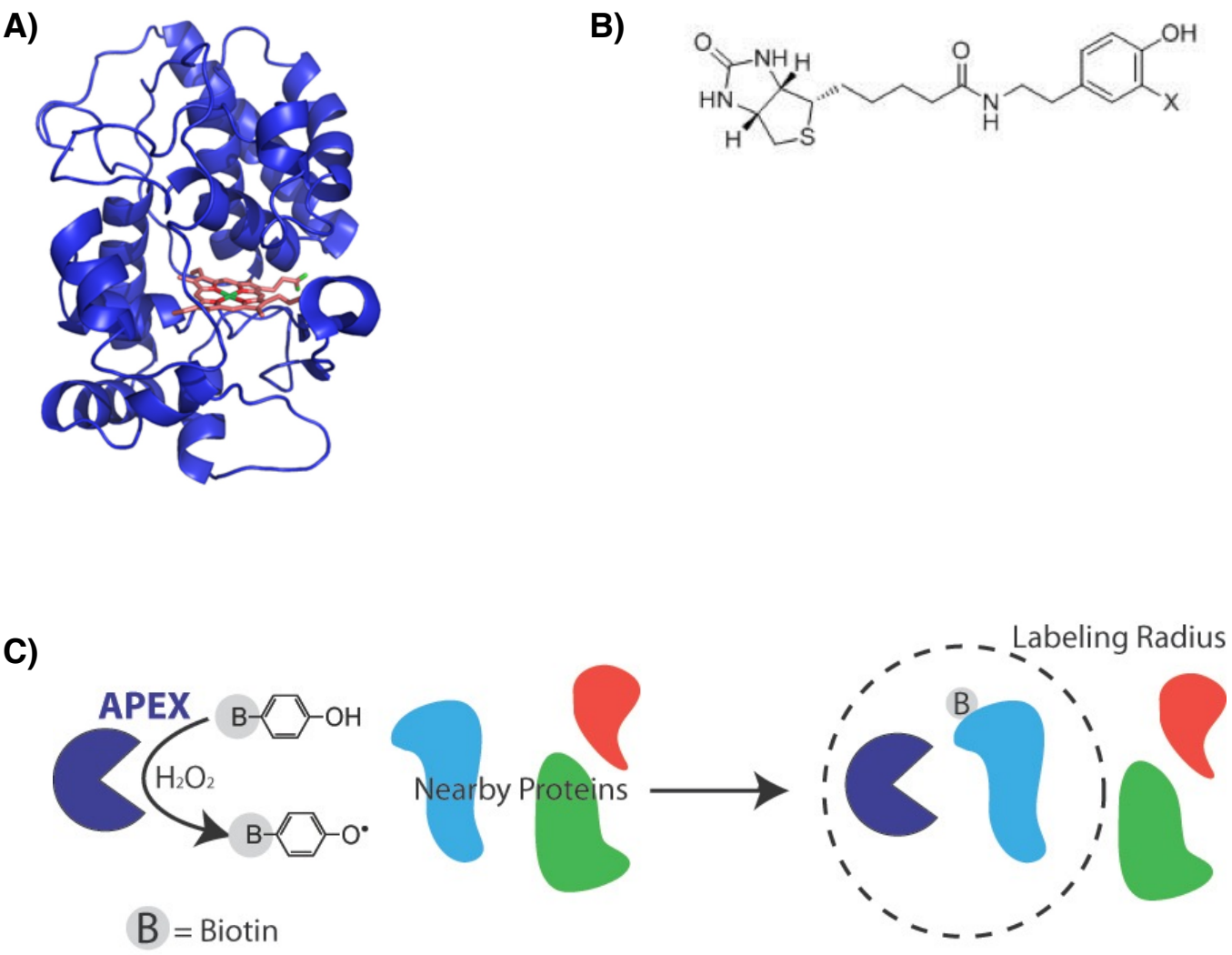


Figure 3.1: The structure and labeling scheme for APEX2. **A)** A crystal structure of soybean ascorbate peroxidase with the heme incorporated into the active site. **B)** A structure of the biotin-phenol compound for covalent tagging (X=H). **C)** A general schematic for APEX labeling. APEX accepts aromatic substrates such as the structure shown in (B). This reactive radical can covalently interact with other exposed aromatic residues. The half-life of the phenoxyl radical is short enough that only proteins in close proximity to APEX will be biotinylated, and can then be enriched using streptavidin.

Results

NFATc2 tolerates fusion to its N- but not C-terminal region

Our first aim was to make a transcription factor fusion to APEX2 that remains transcriptionally active. We fused APEX2 to either the amino or carboxyl termini of NFATc2, separated from the protein by the Human influenza hemagglutinin (HA) epitope tag (see schematics in Figure 3.2). In order to test whether the APEX2 fusions to NFATc2 were transcriptionally active, we utilized the IL-2-Luciferase reporter because, when co-transfected with cJun, it provides a read-out not only on NFAT's ability to bind DNA, but also its ability to interact with other protein factors to mediate high levels of transcription[54].

As shown in Figure 3.2, the normalized Luciferase data show that fusion of APEX2 to the C-terminus of NFATc2 significantly reduced the transcriptional activity of NFATc2. By contrast, transcriptional activation by the amino-terminal fusion was only minimally impaired. Moreover, the level of activation with cJun and APEX2-NFAT was greater than the additive activation of APEX2-NFATc2 and cJun individually, suggesting that the APEX2-NFATc2 is still able not only to bind DNA, but also to interact with cJun in mediating a synergistic activation of IL-2. We chose to use APEX2-HA-NFATc2 for further experimentation to test this as a tool for promiscuous labeling of nuclear proteins in close proximity to NFATc2.

APEX2-NFATc2 labels nuclear protein upon stimulation

To test whether the labeling reaction would be contained to the nucleus, we transfected MDA-MB-231 cells with APEX2-NFATc2. After 20 hours, cells were stimulated with ionomycin and PMA to localize NFATc2 to the nucleus, then labeled 4-7 hours later. The cells were then fixed and stained with DAPI and streptavidin conjugated with Alexafluor555; the later would detect biotinylated protein. Control experiments showed the labeling reaction required all three components of the labeling system: biotin-phenol, H_2O_2, and APEX2. Stimulation with PMA and ionomycin for seven hours was sufficient to ensure that the majority of biotinylation occurred in the nucleus, while unstimulated cells had almost no nuclear staining (Figure 3.3). The localized staining pattern of the neutravidin555 within the nucleus made us confident that the reactivity of the phenol radical with APEX2-NFATc2 is highly specific, and is perhaps detecting where APEX2-NFATc2 is localized to euchromatin.

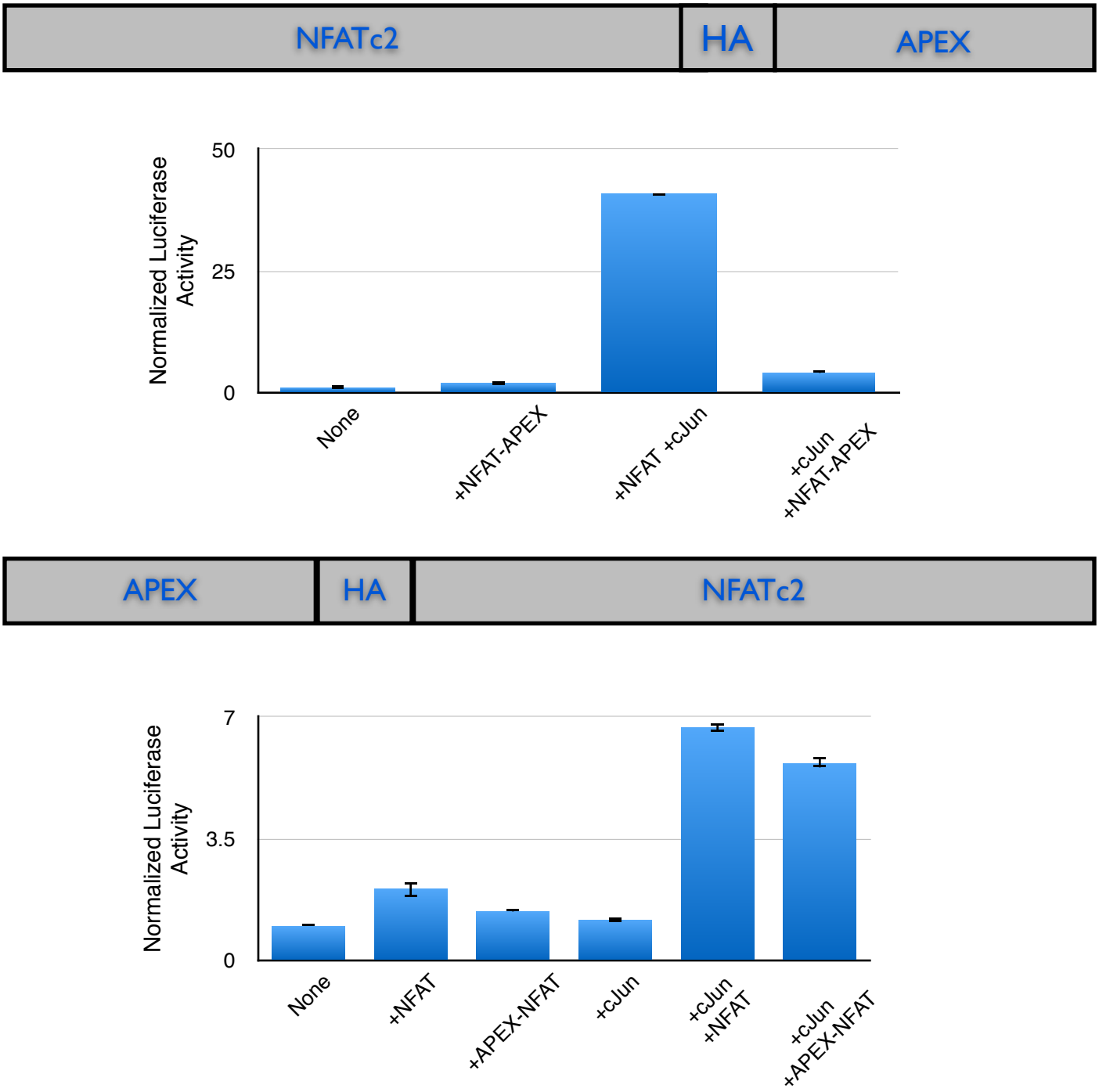


Figure 3.2: Fusion to the N-terminus of NFATc2 allows the protein to remain transcriptionally active. Schematics of the C- and N-terminal fusions to NFATc2 are shown. Cos7 cells were cotransfected with plasmids coding for IL-2-Luciferase and Renilla-null. Relative Luciferase activity was determined by dividing the Firefly Luciferase activity by the Renilla Luciferase activity. The data were normalized to the average activity in cells not over expressing either cJun or NFATc2. Data represent the average of three transfections and error bars represent one standard deviation.

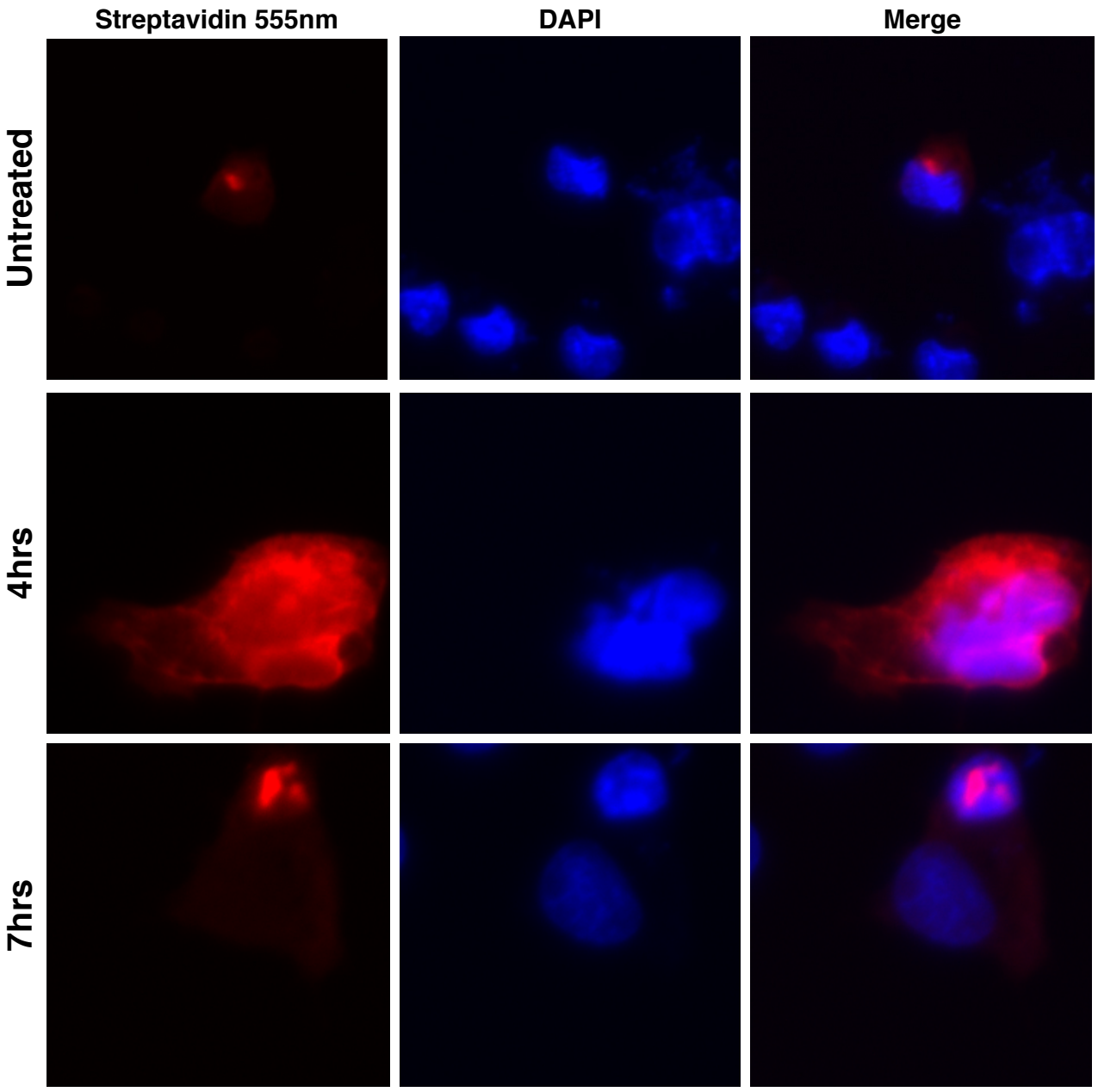


Figure 3.3: APEX2-NFATc2 labels nuclear proteins after stimulation. MDA-MB-231 cells transiently transfected with APEX2-NFATc2 were stimulated with PMA and ionomycin for between 0 and 7 hours, then labeled with biotin-phenol and fixed. Labeled cells were stained with DAPI to mark the nucleus, and with Streptavidin-555 to detect biotinylated proteins.

To test whether NFATc2-specific biotinylated nuclear proteins could be detected from a cellular lysate, cells were transfected with APEX2-NFATc2 or with APEX2-NLS, which contains a nuclear import signal that keeps the protein constitutively in the nucleus. After 20 hours, cells were stimulated with PMA and ionomycin, and APEX labeling was performed seven hours later. After labeling, cells were pelleted and separate nuclear and cytosolic extracts were prepared. Biotinylated proteins in the extracts were detected by Western blotting against biotin. As Figure 3.4A and B show, nuclear fractions of unstimulated cells have a much weaker label intensity across the lane. In cells that were stimulated, their nuclear fractions were highly more enriched for the biotin label, though the difference between those treated with lithium acetate before stimulation and cells that were just stimulated is small, suggesting that the lithium acetate is not required for nuclear labeling. The APEX2-NLS positive control shows the strongest biotin signal, which was expected given the localization pattern of NFAT that we observed when using mApple-NFATc2.

Streptavidin enrichment pulls down protein interacting partners

An important test to ensure that the APEX2 probe is functioning as we expected is to look for labeling of proteins known to interact with NFATc2. To do this, we transfected APEX-NFATc2 or APEX-NLS into MDA-MB-231 cells, and incubated for 20 hours. Cells were then stimulated with PMA and ionomycin for seven hours before labeling. After labeling, the nuclei were isolated and the nuclear extracts were incubated with neutravidin beads. Simultaneous immunoblots of the enriched nuclear extract against NFATc2 and cJun show that both proteins were labeled by APEX2. Additionally, there was a significant amount of self-labeling of APEX2-NFATc2, which is an important positive control (Figure 3.4 C and D). While the low abundance of either protein makes quantitative analysis unreliable, these data give us confidence that we can continue to work toward a quantitative assay.

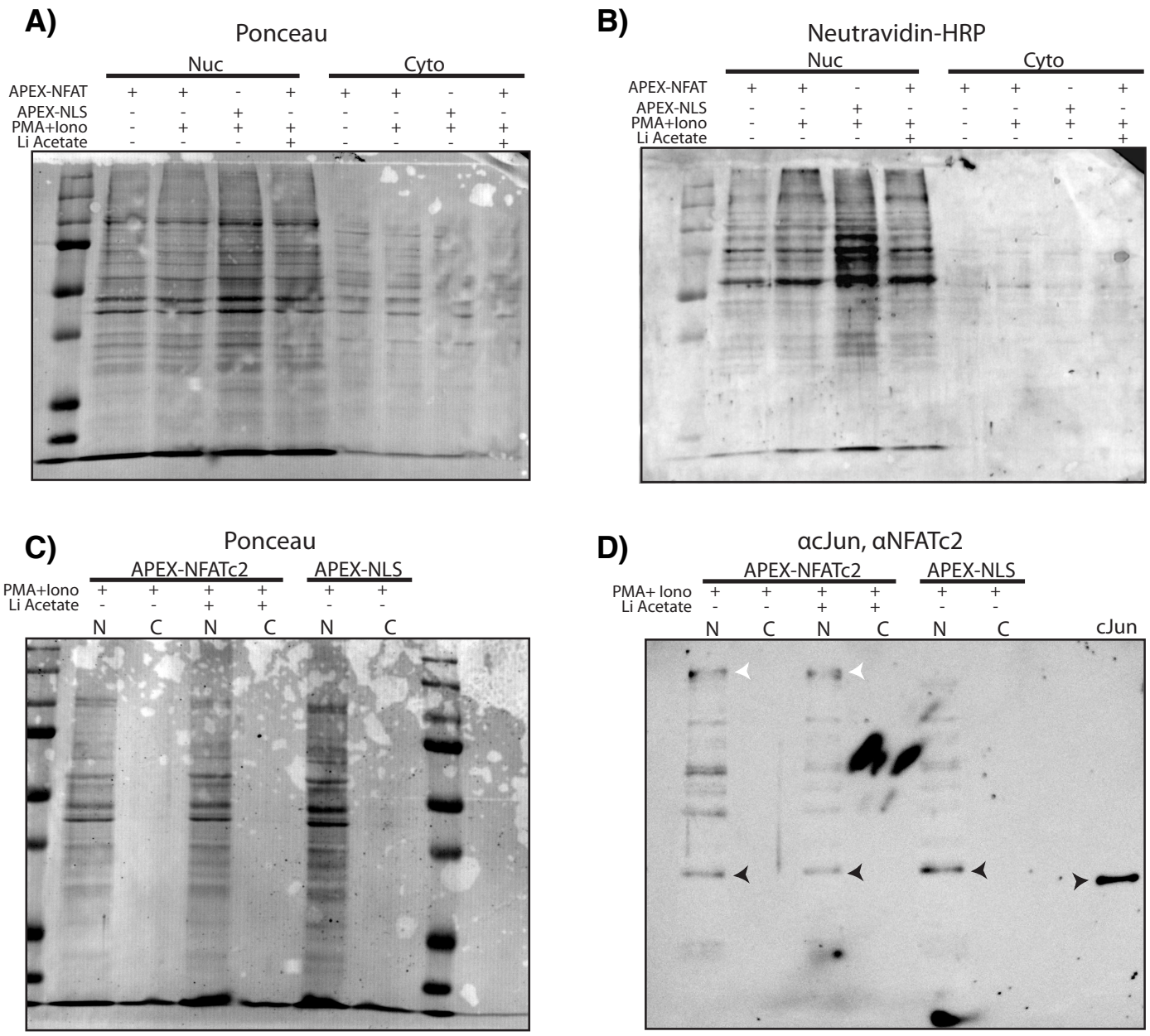


Figure 3.4: APEX labeling tags nuclear proteins. MDA-MB-231 cells were transiently transfected with APEX2-NFATc2 or APEX2-NLS, induced for 7 hours in the presence or absence of lithium acetate, and the cells labeled. Nuclear and cytoplasmic fractions were separated and the labeled proteins visualized by ponceau stain **(A)** and immunoblotting against biotin **(B)**. Stimulated cells were labeled, and the nuclear and cytoplasmic fractions enriched over neutravidin beads. Staining by ponceau shows more protein captured in the APEX-NLS samples **(C)**. Immunobloting against cJun and NFATc2 simultaneously show that APEX2-NFATc2 labels itself (white arrow), as well as labels cJun (black arrow) **(D)**.

Discussion

Understanding a process as complex as the regulation of gene expression in mammals requires a large arsenal of tools. Techniques such as ChIP and its new age derivatives can generate a great amount of data about the genomic occupancy of any DNA binding protein, but it requires a knowledge of what protein to look for and the limits of antibody detection can make it hard to distinguish between specific isoforms or post translational modifications. CoIP, when coupled with mass spectrometry, does not need fore-knowledge of which proteins might interact with the immunoprecipitated protein, but it requires protein interactions that can stand up to numerous washes during the extraction process.

We wanted to see if APEX could be a tool to fill in the gaps left by these two techniques. By fusing APEX to a transcription factor, we showed that we can obtain stable labeling of nuclear proteins which interact in proximity to that transcription factor (Figure 3.3). The ability to detect known interacting partners, like cJun, out of isolated nuclei suggests that the APEX2 labeling may be restricted enough to differentially detect proteins associated with NFATc2. Looking forward, the APEX-NLS construct could be used as a means of background subtraction such that comparing an LC-MS/MS run of the streptavidin enriched nuclear extracts from APEX-NLS and APEX-NFATc2 might yield a list of highly enriched proteins with specificity for interacting with NFATc2. This technique, then, is not limited by the binding kinetics of the activators and coactivators which associate with a promoter and a network of other proteins, but instead is limited only by the throughput of the mass spectrometer and the scientist's ability to subsequently test each of these potential interactions.

We showed that NFATc2 can remain transcriptionally active even when fused to a large protein. If one were to repeat this process with other transcription factors (for example cJun and NF-κB), and determine their nearby protein interactions in the same way, one could use the intersections from their LC-MS/MS data to construct a more extensive picture of the protein species interacting with any given transcription factor. Combined with other data such as ChIP-seq or CoIP, we have the potential to greatly expand our understanding of mechanisms that regulate the combinatorial control of transcription.

Materials and Methods

Cell culture

Human mammary adenocarcinoma cells (MDA-MB-231) were cultured in Libovitz's L-15 medium supplemented with 10% (v/v) fetal bovine serum, 2 mM L-glutamine, 100 U/mL penicillin and 100 μg/mL streptomycin. Cells were trypsinized with 0.25% trypsin and subcultured roughly every four days, and maintained at 37 °C and 0% CO_2 in a humid incubator. Genetically encoded reporters were transfected using TransIT-BrCa (Mirus Bio) or electroporated using the Neon transfection system (Invitrogen) according to the manufacturer's instructions.

Cercopithecus aethiops kidney cells (Cos7) were cultured in DMEM supplemented with 10% (v/v) fetal bovine serum, 2 mM L-glutamine, 100 U/mL penicillin and 100 μg/mL streptomycin. Cells were trypsinized with 0.25% trypsin and subcultured roughly every two days, and maintained at 37 °C and 5% CO_2 in a humid incubator. Genetically encoded reporters were transfected using TransIT-2020 (Mirus Bio) according to the manufacturer's instructions.

Plasmid construction

The construction of pcDNA 3.1(-) HA-NFATc2 is described above. To make APEX2-HA-NFATc2, the APEX2 insert from pcDNA 3.1(+) APEX2-NLS (a generous gift from Alice Y. Ting, Cambridge, MA) was amplified using primers containing XbaI and XhoI restriction sites. pcDNA 3.1(-)-HA-NFATc2 was digested with XbaI and XhoI, followed by CIP treatment. The vector and insert were ligated before transformation. NFATc2-HA-APEX2 was constructed by amplifying APEX2 with NotI and EcoRI restriction sites, and NFATc2 with XbaI and XhoI restriction sites. The two inserts were sequentially ligated into the digested APEX2-HA-NFATc2 plasmid. All constructs were verified by sequencing.

Biotin-phenol synthesis

Synthesis of the biotin-phenol compound was performed by Will Hartwig (Boulder, CO), and has been described[63]. The compound was purified through several rounds of azeotrope formation by pyrimidine/toluene fractional distillation. The distillate was then loaded onto a column and the resulting crude dissolved in methanol and removed under vacuum with standard dry silica gel. The compound was then dry loaded and purified once more through flash chromatography in

9:1 chloroform:methanol and crystalized. The material was washed with chloroform and dried under vacuum to yield a pure powder. The biotin-phenol was dissolved in tissue culture grade DMSO to a concentration of 500 mM, aliquoted and stored at -80 °C.

APEX2 labeling reaction

Cells transfected with the APEX2-NLS or APEX2-HA-NFATc2 plasmids were grown for 20 hours. 30 minutes prior to labeling, cell medium was exchanged with warm PBS containing 500 μM biotin-phenol. To label, 1 mM H_2O_2 was added to the cells, and labeling was allowed to proceed for 1 minute. After 1 minute, a quencher solution (10 mM sodium azide, 10 mM sodium ascorbate, 5 mM 6-hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid (Trolox) (Sigma Aldrich) was added and allowed to incubate on ice for one minute. At this point, cells were either fixed in a fixative (10 mM sodium azide, 10 mM sodium ascorbate, 5 mM Trolox, 4% paraformaldehyde), or collected for protein extraction.

Immunostaining

Cells to be immunostained were fixed in 4% paraformaldehyde for 15 minutes at room temperature. After 15 minutes, the fixative was removed and the cells washed 3 times in PBS. To permeabilize, an appropriate amount of -20 °C methanol was added to the cells, and the plate incubated for 5 minutes at -20 °C. The methanol was removed and the cells thoroughly washed with PBS before blocking for 1 hour at room temperature using blocking buffer (3% bovine serum albumin or 1% casein in PBS). To stain against biotinylated protein, AlexaFluor555-conjugated streptavidin was added to blocking buffer to a final concentration of 500 ng/mL, and added to the cells for 30-60 minutes. After incubation, the cells were washed three times with PBS+ 0.1% Tween20, then stained with a 300 nM solution of DAPI (Life Technologies) in PBS+ 0.1% Tween20. Finally the cells were washed another three times with PBS+ 0.1% Tween20, followed by two washed in PBS. Cells were imaged using an EVOS FL Cell Imaging System (Life Technologies).

Protein extraction and streptavidin enrichment

Labeled cells were scraped from the bottom of the plate, then transferred to 1.5 ml eppendorf tubes and centrifuged at 1100x g for 5 minutes. The supernatant was removed, and the cells suspended in 200 μL NROQ buffer (10 mM Tris pH 7.9, 10 mM NaCl, 4 mM $MgCl_2$, 10 mM

sodium azide, 10 mM sodium ascorbate, 5 mM Trolox, 1 mM PMSF, 1x Protease Inhibitor Cocktail (Roche)) with 0.05% NP-40. Cells were allowed to lyse gently on ice for 10 minutes, then spun at 1100x g for 5 minutes. The supernatant was removed, and the nuclei washed once in 100 µL NROQ with 0.05%. After washing, the isolated nuclei were suspended in NROQ with 0.15% NP-40 and vortexed in pulses for 15 seconds, then incubated on ice for another 10 minutes.

For neutravidin enrichment, 75 µL of 50/50 bead slurry per reaction was centrifuged at 1100x g for 5 minutes. The beads were suspended in NROQ buffer and allowed to equilibrate for 5 minutes, before centrifugation at 1100x g. After 3 rounds of washes, the supernatant was removed from the equilibrated beads, and cytosolic or nuclear fractions loaded onto them. Binding was allowed to proceed for 1 hour while nutating, then the beads were spun at 1100x g and the supernatant removed. To remove nonspecific binders, beads were suspended in 250 µL NROQ and allowed to nutate for 30 minutes before centrifugation. After the supernatant was removed, beads were suspended in 50 µL protein loading buffer (120 mM Tris HCl pH 6.8, 4% SDS, 0.02% bromophenol blue, 20% glycerol) and incubated for 3 minutes at 95 °C, vortexing occasionally. Samples were resolved with SDS-polyacrylamide gel electrophoresis (SDS-PAGE), and visualized via silver stain, coomassie stain, or immunoblotting.

Luciferase assays

250 ng of each protein expression vector, or blank vector, were transfected into Cos7 cells at 60% confluence in a 12-well plate using TransIT-2020 (Mirus Bio) as per the manufacturers instructions. 20 hours post transfection, cells were stimulated with 20 ng/mL PMA and 1 µM ionomycin for 7 hours. After 7 hours, cells were harvested in 200 µL 1x Passive Lysis Buffer (Promega) for 15 minutes at room temperature. Firefly and Renilla luciferase values were determined using the Dual-Luciferase Kit (Promega), and the Firefly values normalized to the Renilla values in each well.

Chapter 4: Characterization of an aptamer that selectively binds cJun/cJun homodimers

Introduction

The AP-1 family of proteins is a heterogeneous group of transcription factors containing a conserved basic leucine zipper (bZip) domain, and which recognize and bind the consensus element 5'-TGA(C/G)TCG-3', also known as the TPA response element[27]. AP-1 proteins dimerize through their bZip domains, and bind DNA to drive dimer-specific transcriptional activation[28]. While the roles of individual AP-1 proteins are generally well understood in aggregate, the contributions of specific dimer pairs in regulating distinct transcriptional programs remain mostly unclear.

Recent evidence for functions unique to cJun/cJun homodimers has emerged[53], underscoring the need to generate methods to study these individual dimers in the cellular context. Current methods to address the roles of individual AP-1 proteins, such as the creation of knock-down or knock-out cell lines, do not allow dimer specificity to be addressed. cJun knock-down, for example, not only removes cJun/cJun homodimers from the cell, but also all other cJun dimer combinations.

To this end, we employed systematic evolution of ligands by exponential enrichment (SELEX) to identify a family of single-stranded DNA aptamers which bind cJun/cJun homodimers tightly and specifically, with the ultimate goal of using these aptamers in cells. SELEX is a technique that takes advantage of the energetically favorable tertiary structures that single-stranded oligonucleotides form[64]. Aptamers were selected from an initial library of approximately 10^{15} oligos 78 nucleotides in length. Each DNA sequence consisted of a 40 nucleotide randomized region flanked by constant regions on their 3'- and 5'-ends. These aptamers were introduced to DNA-bound cJun/cJun homodimers (Figure 4.1). After washing, nucleotide sequences that associated with the DNA-bound cJun/cJun were amplified using primers against their constant regions. The ssDNA SELEX pool was regenerated by isolating the appropriate strand from the double stranded PCR product, and reintroduced to the target under more stringent wash conditions. After 6 rounds, including a negative selection in rounds 5 and 6 against off-target binders, 46 clones were sequenced[65].

Here, we screened and characterized a number of these aptamers for their ability to bind directly to cJun/cJun. We identified an aptamer which binds cJun/cJun homodimers with greater than 100-fold specificity over cJun/cFos heterodimers, and used a variety of biochemical techniques to better understand the interaction between cJun/cJun and the aptamer.

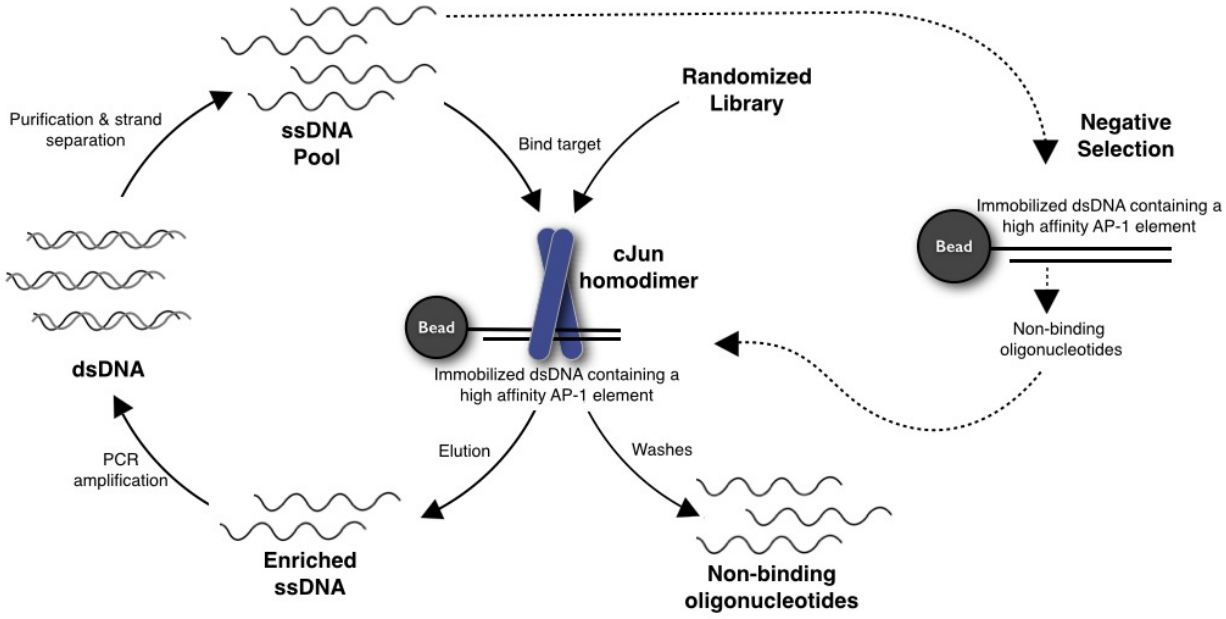


Figure 4.1: A schematic of the selection process for identifying an aptamer against cJun/cJun homodimers. Purified cJun/cJun was immobilized by binding an AP-1 element. The randomized nucleotide library was allowed to fold into energetically favorable conformations, then introduced to the purified protein. Six rounds of SELEX were performed, with the last two rounds including a negative selection against off-target binders.

Results

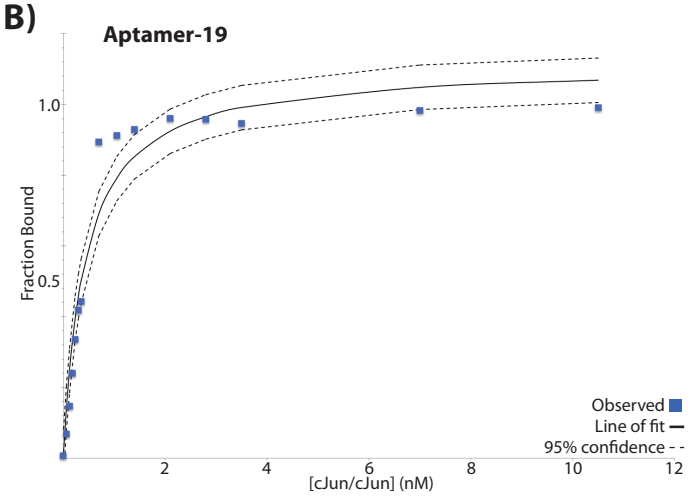
Binding kinetics suggest aptamers have high affinity and selectivity towards cJun/cJun homodimers

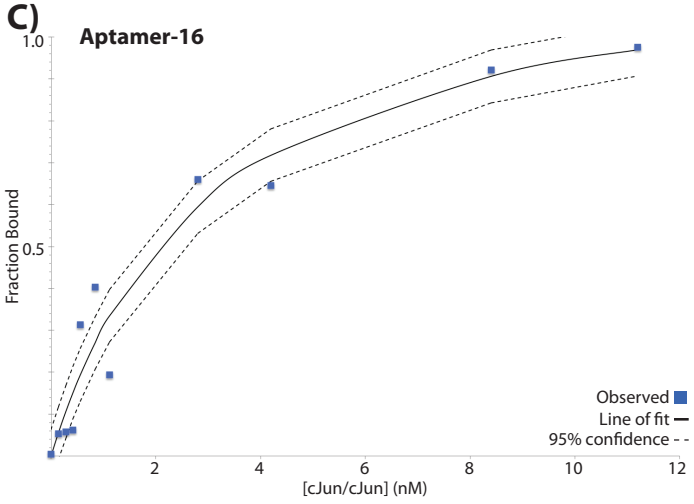
All 46 sequences were different but contained over-represented motifs in the 5' and 3' ends of the randomized region, suggesting similar responses to the selective pressure from SELEX; aptamers 16 and 19 were chosen for further study. Sequence alignments of aptamer-16 and aptamer-19 illustrate the common motifs at the 5' and 3' ends of their randomized regions (Figure 4.2A). Electrophoretic Mobility Shift Assays (EMSAs) with ^{32}P-labeled aptamer-16 and aptamer-19 show that both bound cJun/cJun homodimers with low- to sub-nanomolar affinity (K_d of 2.6 nM and 0.4 nM, respectively) (Figure 4.2 B-C). As a point of comparison, we also measured the affinity of cJun/cJun for its DNA recognition element, using an AP-1 decoy. The AP-1 decoy, a single hairpin containing the consensus AP-1 recognition element, bound cJun/cJun homodimers with a K_d of 7.6 nM, consistent with previously reported data[66] (Figure 4.2 D). Aptamer-19 showed an order of magnitude tighter binding than the AP-1 decoy, which motivated us to select this aptamer for further study.

In order to address the selectivity of aptamer-19, we wanted to compare the aptamer's affinity for binding cJun/cJun homodimers to that for binding cJun/cFos heterodimers, another well-established dimer combination in the AP-1 family which is thought to activate a different transcriptional program than cJun/cJun homodimers. In order to do this, we cloned and expressed 6xHis-cFos in *E.coli* and purified the recombinant protein using a Nickel-NTA column. Purified cJun and cFos were then combined in stoichiometric amounts and dialyzed. EMSAs using radio-labeled aptamer-19 showed a roughly 200-fold lower affinity for cJun/cFos heterodimers than for cJun/cJun homodimers (Figure 4.3 A), despite the fact that cJun/cFos heterodimers have a higher affinity for the consensus AP-1 decoy. The low affinity for cJun/cFos heterodimers suggests that aptamer-19 has a high degree of specificity for cJun/cJun homodimers. Measurements of the off-rate for the cJun/cJun-aptamer-19 complex reveals a relatively stable aptamer-protein complex, with a half-life of approximately 16 minutes (Figure 4.3 B). Using the value of the dissociation rate constant and the K_d we could calculate an observed association constant of 4.1×10^6 M^{-1} s^{-1}.

A)

```
Apt-16:gggagatcacttacggcaccCGGAGGgtcgttac--ccactgtttgttg--gtataaagcctg--ccaaggctcgggacagcg
Apt-19:gggagatcacttacggcacc-gtatagtcgt-acatgaa-cgagtgtgagtgt-t-aagcctttgccaaggctcgggacagcg
```





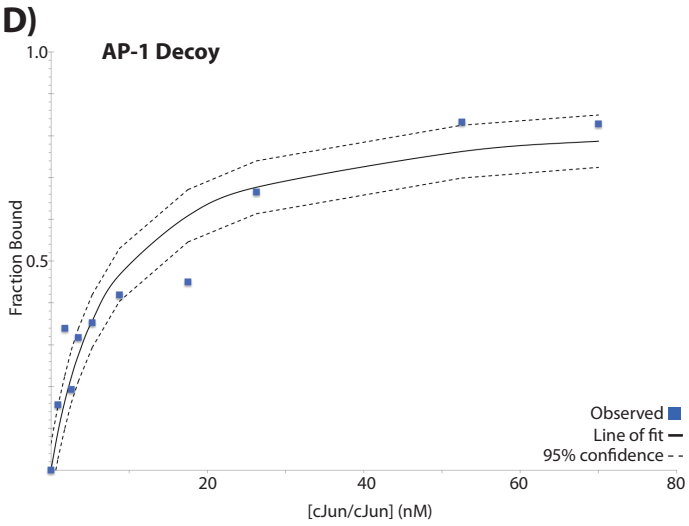


Figure 4.2: Selection yielded aptamers with a higher affinity for cJun/cJun homodimers than the consensus DNA binding sequence. **A)** Alignment of aptamer-16 and aptamer-19 shows two conserved motifs in the randomized region (indicated by a grey box). 5' and 3' constant regions are shown in bold. **B-D)** Aptamer-19 and aptamer-16 bind cJun/cJun homodimers with a higher affinity than a single-stranded DNA hairpin containing the AP-1 consensus binding sequence. Aptamer-19 binds cJun/cJun homodimers with sub-nanomolar affinity (K_d=0.4nM), whereas aptamer-16 binds and the AP-1 decoy bind slightly less tightly (K_d=2.6nM and K_d=7.6nM respectively).

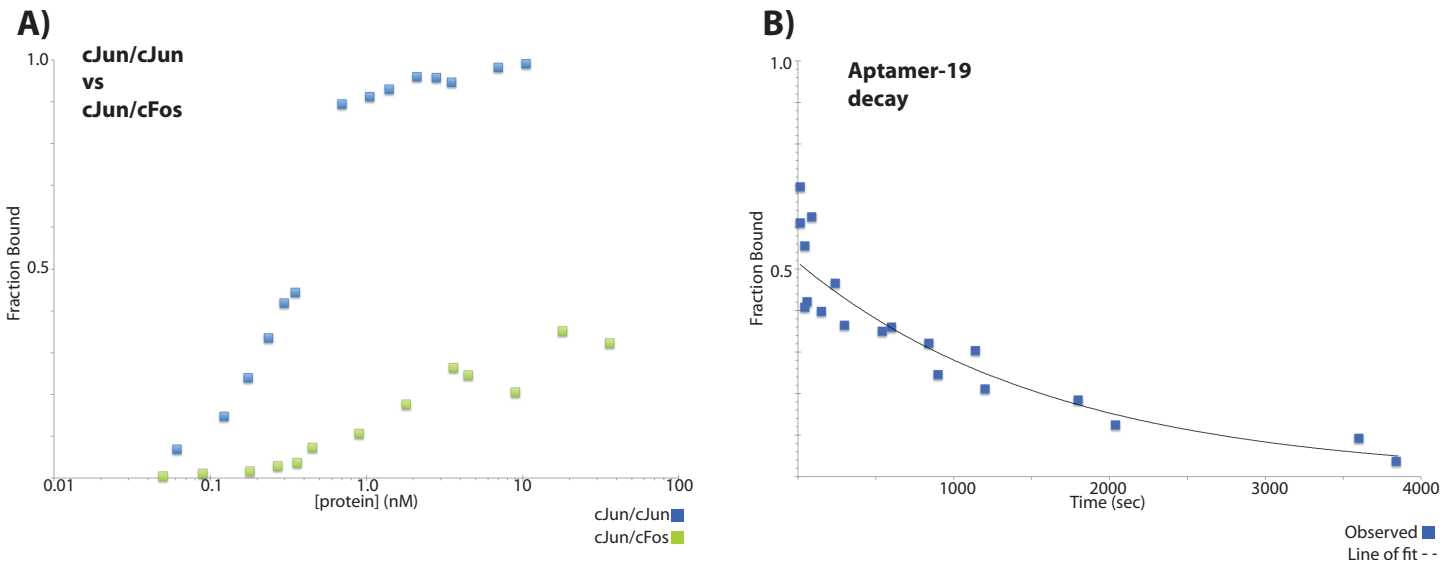


Figure 4.3: Aptamer-19 selectively and stably binds cJun/cJun. **A)** Aptamer-19 has a much higher affinity for cJun/cJun homodimers than for cJun/cFos heterodimers. EMSAs titrating a broad range of protein concentrations into ^{32}P-labeled aptamer-19 were quantitated and plotted. Aptamer-19 binds cJun/cJun homodimers with sub-nanomolar affinity (K_d=0.4nM, shown in figure 4.2B), whereas it binds cJun/cFos heterodimer approximately 200-fold weaker (K_d=85nM). **B)** The off-rate of aptamer-19 was determined through EMSA by incubating ^{32}P-labeled aptamer with cJun/cJun homodimers and allowing complexes to form. 1000-fold excess unlabeled aptamer was added, and the reaction loaded onto the gel at increasing time points. Aptamer-19 has a dissociation rate k_d= 7.1x10^{-4} s^{-1}, and a half-life of ~16 minutes.

S1 Nuclease reveals regions of the aptamer that are single stranded

The ability of aptamer-19 to bind cJun/cJun with high affinity and specificity compared to other AP-1 dimer combinations inspired us to investigate some of the structural features of the aptamer. When given the nucleotide sequence of aptamer-19, structure prediction software such as Mfold returned a large number of possible secondary structures based on thermodynamic considerations[67, 68]; experimental information would be required to assess the secondary structure of the aptamer.

In order to experimentally probe the secondary structure of aptamer-19, we employed S1 nuclease, which cleaves specifically at single-stranded regions of DNA. ^{32}P-labeled aptamer-19 was allowed to fold into its most thermodynamically stable conformation, and then was incubated on ice with increasing amounts of S1 nuclease. After resolution of digested products on a denaturing gel, 16 nucleotide positions had a signal at least two times greater than the undigested lane (Figure 4.4 A). These data were used as constraints in the structure prediction programs, which yielded a single most likely candidate for the structure of the aptamer (Figure 4.5).

We noticed that the third stem loop of the predicted structure contains a recognition site for the restriction endonuclease CviKI-1, a blunt-end endonuclease which recognizes the double-stranded sequence 5'-RGCY-3'. To test whether our structure prediction aligned with experimental evidence, we incubated the folded aptamer with CviKI-1. As the secondary structure predicted, CviKI-1 cleaved the aptamer, with the major product being 54 nucleotides long, and a minor product which was 67 nucleotides; this latter product is most likely the result of an incomplete cleavage event (Figure 4.4 A). Additionally, we made a modified aptamer in which eleven nucleotides on the 5' end and four nucleotides from the 3' end (both of which were predicted to be single-stranded) were removed. This truncated aptamer-19 (12-74) was digested with S1 nuclease and CviKI-1 in an identical manner to the full-length aptamer and the fragments resolved on a denaturing gel. Aptamer-19 (12-74) showed a pattern of digestion very similar to that of the full-length aptamer (Figure 4.4 B). Cleavage with CviKI-1 yielded a major product which was 44 nucleotides in length and a minor product 55 nucleotides in length, as the secondary structure would predict. The smaller aptamer also allowed better resolution of the

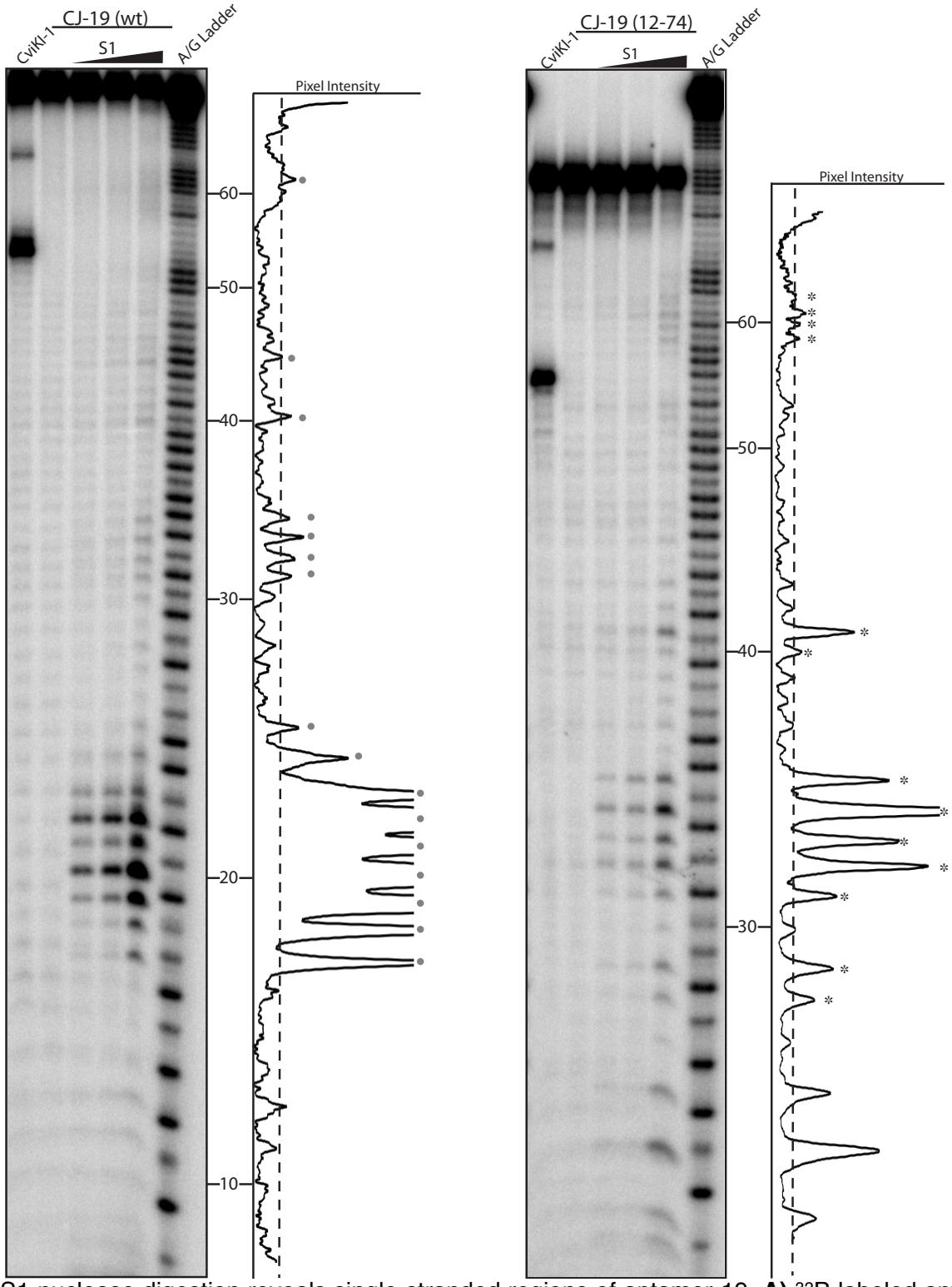


Figure 4.4: S1 nuclease digestion reveals single-stranded regions of aptamer-19. **A)** ^{32}P-labeled aptamer-19 was incubated with the restriction endonuclease CviKI-1, or with increasing amounts of S1 nuclease. The resulting fragments were resolved on a denaturing gel. The plot next to the gel represents the pixel intensity of lane 5, the dashed line is the cut off used to determine single stranded nucleotides, and position numbers refer to the full-length aptamer. Digestion with CviKI-1 yields a 54 nt major product and a 67 nt minor product, consistent with our predictions. S1 cleavage of the full-length aptamer reveals four major single-stranded regions. **B)** To achieve better resolution at the 3' end of the aptamer, aptamer-19 (12-74) was digested exactly as before. Consistent with the full length aptamer, when digested with CviKI-1 aptamer-19(12-74) has a 44 nt major product and a 55 nt minor product. Additionally, S1-nuclease digestion of aptamer-19(12-74) reveals a single-stranded region between nucleotides 47 and 51, corresponding to the unresolved peaks towards the top of **(A)**.

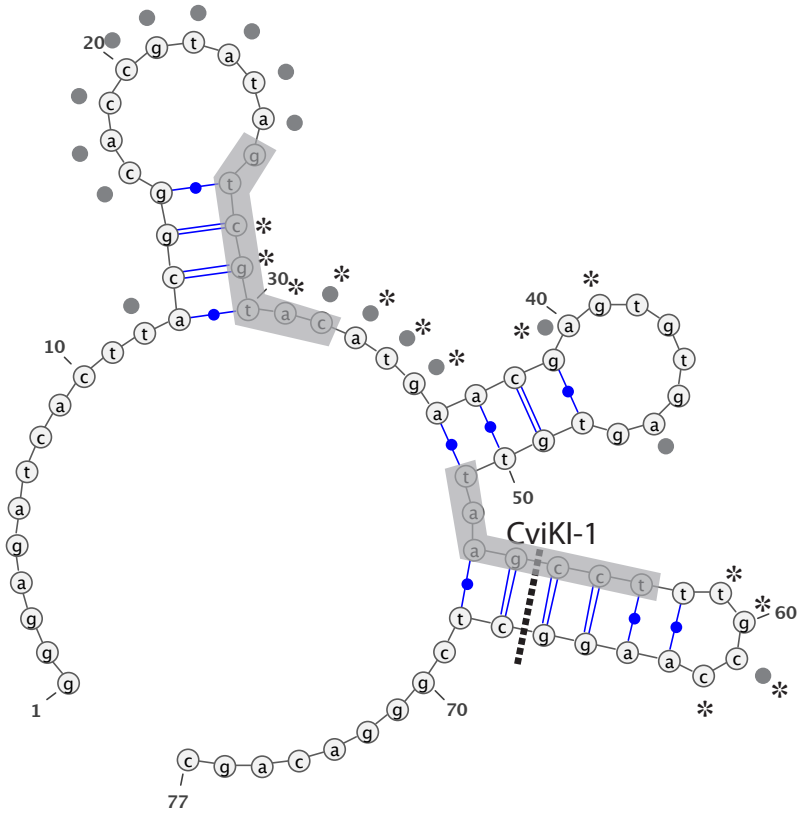


Figure 4.5: Data from S1-nuclease digestion reveals a probable secondary structure of aptamer-19. Grey dots indicate single-stranded regions revealed by S1 digestion of the full-length aptamer, while asterisks denote single-stranded regions in the 12-74 nt truncation. The position of the CviKI-1 digestion site is marked by the dashed line. Using these data as parameters, we were able to predict a secondary structure. The two conserved motifs between aptamer-16 and aptamer-19 are highlighted in grey.

upper bands on the gel, allowing us to detect individual cleavage events by S1 nuclease in the 47-51 nucleotide region, which supports the secondary structure model.

DNase I footprinting suggests that aptamer-19 may target unique elements in the activation domain of cJun/cJun homodimers

Given the higher affinity of aptamer-19 for cJun/cJun homodimers, we asked whether this interaction was caused by an interaction within the domains unique to cJun/cJun homodimers. In order to probe this, we separately expressed the bZip domain of cJun, which is highly conserved within AP-1 family proteins. We evaluated the ability of full length cJun/cJun to protect aptamer-19 from cleavage by DNase I compared to the bZip region of cJun. DNase I cleaves DNA in a non-sequence-specific manner, but given the inherent bulk of the enzyme, it cannot access nucleotides which are already associated with protein. This technique, then, allows us to determine which regions of the aptamer are actively associated with protein.

Incubation of aptamer-19 with DNase I revealed seven regions which are cleaved with high frequency, most likely owing to either the preference of DNase I for double stranded DNA over single-stranded, or the solvent accessibility of certain regions in the aptamer (Figure 4.6.A, B). When cJun/cJun homodimers were added to the reaction, cleavage of the aptamer was greatly reduced in all but one position on the aptamer, suggesting that multiple regions of the aptamer are involved in interactions with the proteins.

When incubated with cJun-bZip, even at high concentrations, there was very little protection due to protein binding. Unfortunately, experiments with the AP-1 decoy and cJun-bZip do not show the protection that we would have expected (data not shown), so the lack of protection is not conclusive evidence. Still, these data favor a model in which the aptamer binds some unique feature on the cJun/cJun homodimer that the bZip dimers alone lack.

Systematic truncations of aptamer-19 suggest a minimal binding domain required for cJun/cJun binding

Motivated by the idea that minimizing the size of the aptamer to only the regions required for tightly binding cJun/cJun would reduce off-target effects when used in cells, we sought to identify the minimal binding domain of aptamer-19. With a proposed secondary structure, as well as DNase I data in hand, we attempted to identify the minimal region by making systematic

truncations from both ends of the aptamer. Six truncations were chosen that could reduce the size of the aptamer from 78 nucleotides to 31 nucleotides. As we anticipated, removal of the 5' single-stranded region had a negligible effect on binding affinity, and in fact reduced the occurrence of multimer aggregation in electrophoretic mobility shift assays (EMSAs) (Figure 4.7). Surprisingly, destruction of the first stem loop by removing the first 22 nucleotides from the 5' end reduced binding affinity of aptamer-19 by only approximately 10-fold. Any further truncation completely obliterated cJun/cJun homodimer binding over the range shown.

Removal of nucleotides from the 3' end showed a much higher dependance on the 3' single-stranded region. The first four nucleotides did not contribute to binding affinity, but removal of the next 8 resulted in a significant decrease in binding affinity. Unsurprisingly, a truncation which destroyed the third stem loop completely abolished protein binding. Given these data, we chose to use the 12-74 truncation as a possible tool for in vivo experimentation.

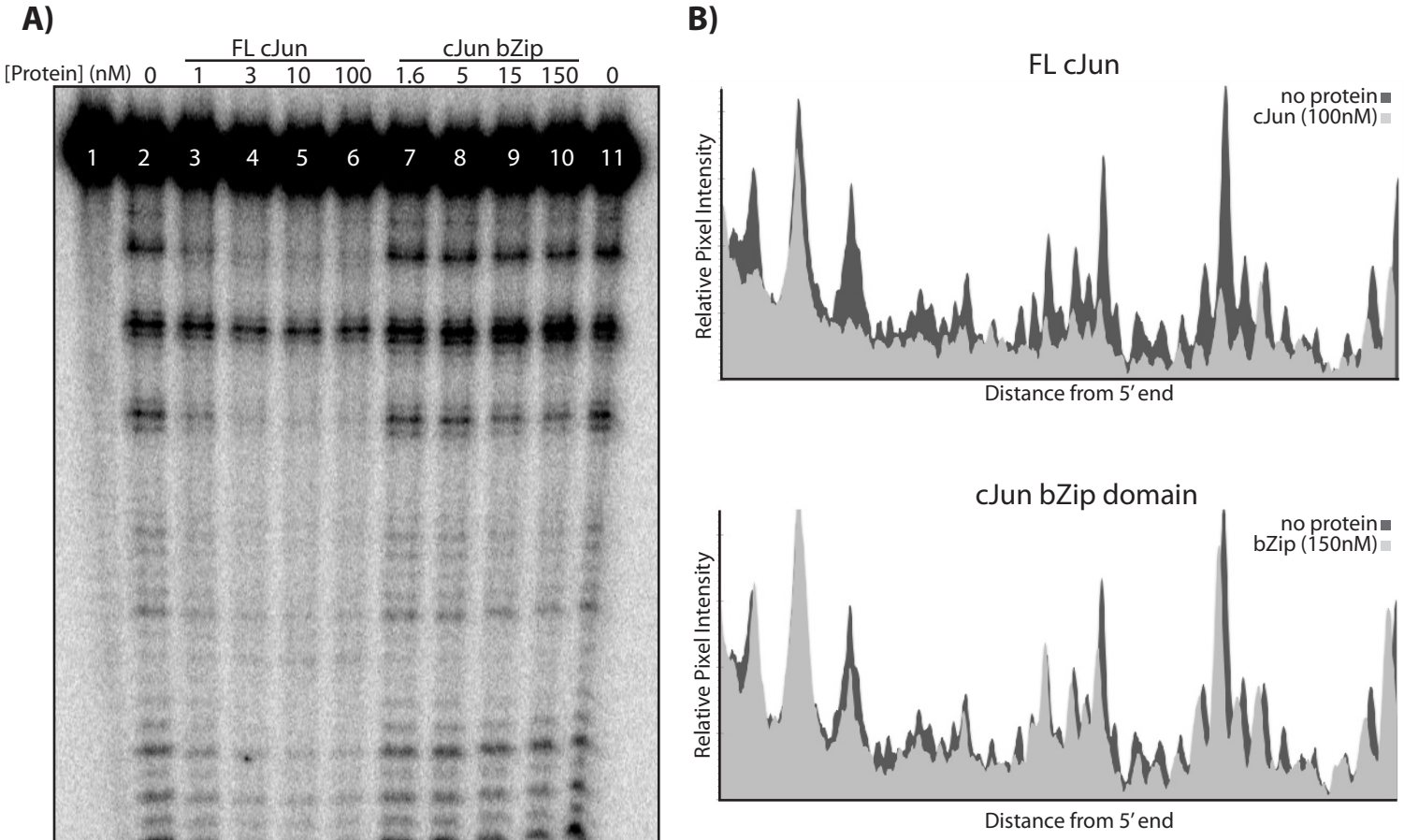


Figure 4.6: A) DNase footprinting of aptamer-19 reveals a broad range of contacts with cJun/cJun homodimers. ^{32}P-labeled aptamer-19 was incubated with increasing amounts of either cJun/cJun homodimers (lanes 3-6) or with dimers of the basic leucine zipper (bZip) domain (lanes 7-10) and compared to aptamer-19 with no protein added (Lanes 2 and 11). Complexes were allowed to form for 20 minutes, then DNase1 was added for 1 minute. Digested fragments were precipitated in ethanol and then run on an 11% polyacrylamide gel. **B)** Plots comparing pixel intensity of bands with or without protein are shown above. Aptamer-19 is broadly protected, even at relatively low amounts of cJun/cJun homodimers, whereas it is not protected by the bZip domain.

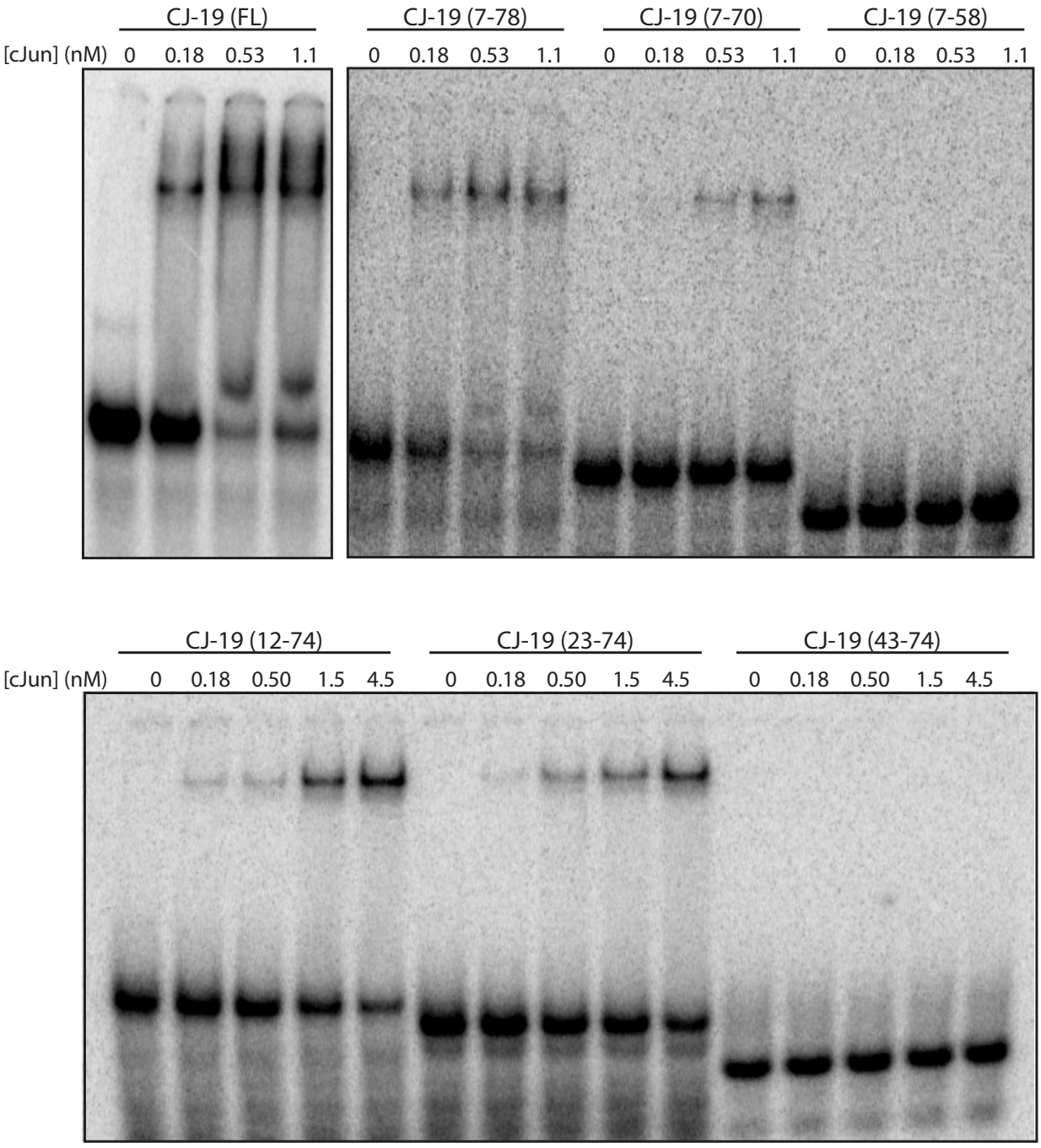


Figure 4.7: Systematic truncations of aptamer-19 reveal a minimal domain required for high-affinity binding. Truncations were made from both the 3' and 5' ends of aptamer-19. Each truncation was ^{32}P-labeled, the incubated with cJun/cJun homodimers for 20 minutes before being loaded into an EMSA. These data show that nucleotides out to position 74 are required on the 3' end. The first 12 nucleotides can be removed from the 5' end with minimal loss of binding affinity (K_d= 0.4nM), and binding is still possible with the first 23nt removed, though some affinity is lost.

Discussion

Aptamer-19 binds cJun/cJun homodimers with surprising affinity

We have reported here the characterization of an aptamer which binds cJun/cJun homodimers selectively and with high affinity. Given the pivotal roles of certain AP-1 proteins in the growth and proliferation of multiple forms of cancer, developing a probe which can potentially differentiate cJun/cJun from the many possible permutations of dimers that can exist within the AP-1 family could have a large impact in understanding exactly what role cJun/cJun plays in the cell.

SELEX is a process which is designed to pull out families of aptamers that have high affinity and specificity for a particular factor of interest. Given that the highly basic DNA binding domain of cJun/cJun was masked with the canonical AP-1 sequence throughout the selection process, we did not expect that the aptamer would be able to succeed in surpassing the binding affinity of the native protein to its target sequence. Such a large binding affinity suggests that a low concentration of aptamer would be sufficient to bind cJun/cJun homodimers in the cellular context.

Footprinting data underscore the importance of aptamer structure

The process of SELEX requires constant regions flanking the randomized nucleotide region to be able to amplify and identify the desirable aptamers throughout the selection process. Our experiments with S1 cleavage and DNase I footprinting, as well as analysis of sequence homology between binding aptamers, suggest that tightly binding aptamers contained regions on both 5' and 3' ends which base pair with regions in the constant region of the full-length aptamer. Through truncations of the aptamer, we found that some nucleotides within the constant regions were required for tight binding. Indeed, even some nucleotides that were predicted to be single-stranded were crucial for the function of the aptamer. Without the ability to achieve a high-resolution crystal structure of aptamer-19, we can only speculate as to its tertiary structure or the specific roles of any one of the nucleotides. It is clear, however, that a different choice of constant region would have an effect on the sequence and structure of the aptamers that succeeded in the selection process. This only serves to highlight the power of the SELEX process and underscores its flexibility as a tool for a wide variety of biological problems.

Aptamer-19 (12-74) has the potential to be a powerful tool

Having developed a tool, which we have shown binds cJun/cJun homodimers specifically over cJun/cFos, we now look forward to its implementation as a tool in understanding the role of cJun/cJun homodimers in regulating gene expression. Some instances of dimer-specific AP-1 functions have already been elucidated, either using *in vitro* characterization or carefully targeted knock-downs of certain AP-1 proteins. Unfortunately the scope of these findings are relatively narrow.

With the ability to target a specific dimer combination, one can imagine that an entire realm of possibilities opens up. RNA-seq or Pol II ChIP-seq, in addition to a cJun ChIP-seq, using cells treated with aptamer-19 could reveal the loci on the genome at which cJun/cJun homodimers act specifically to regulate the expression of protein-coding messages. Further experimentation and modification of the existing aptamers could result in a tool that can specifically target cJun/cJun homodimers for export from the nucleus and subsequent degradation. Aptamers specific for other AP-1 dimer combinations could be developed. This aptamer and the SELEX approach are both valuable tools in understanding the role and scope of cJun/cJun homodimers as AP-1 transcriptional activators.

Methods and Materials

Plasmid preparation

pET-cJun, pET-cJun(c269s), and pET-6xHis-cFos have been described previously[60, 65]. pET-6xHis-cFos(c154s) was constructed using QuickChange site directed mutagenesis in order to limit protein aggregation in EMSAs.

Protein purification

cJun/cJun homodimers were recombinantly expressed and purified as described previously. To make cJun/cFos heterodimers, cultures were co-transformed with pET-6xHis-cFos and pSBET and grown in the presence of 100 μg/mL ampicillin and 50 μg/mL kanamycin in Luria-Bertani broth at 37 °C. When an optical density at 600 nm reached 0.4, cultures were induced with 0.5 mM isopropylthio-β-D-galactoside. After 2 hours, cells were harvested and the cell pellet resuspended in a buffer containing 20 mM Tris (pH 7.9), 1 mM EDTA, 0.1 M NaCl, 1 mM

DTT, 0.2 mM phenylmethylsulfonyl fluoride (PMSF) and sonicated four times for 15 seconds. Samples were centrifuged for 30 minutes at 15,000 rpm and 4 °C. The supernatant was discarded, and the pellet containing 6xHis-cFos was resuspended in 10 mL of 20 mM Tris (pH 7.9), 1 mM EDTA, 0.1 M NaCl, 5 mM DTT, 0.2 mM PMSF and sonicated two times for 30 seconds. The pellet was washed three times by resuspension in 10 mL of 20 mM Tris (pH 7.9), 1 mM EDTA, 0.1 M NaCl, 5 mM DTT, 0.2 mM PMSF. The pellet from the last wash was resuspended in 10 mL of buffer A (20 mM Tris (pH 7.9), 1 mM EDTA, 5 mM DTT, 8 M Urea, 0.1 M NaCl) containing 20 mM imidazole. Soluble material was loaded onto a Ni-NTA agarose column (Qiagen) and washed with 5 column-volumes of buffer A containing 20 mM imidazole, followed by 5 column-volumes of buffer A containing 40 mM imidazole. Bound protein was eluted from the column using buffer A containing 500 mM imidazole. Purified 6xHis-cFos was mixed with an equal molar amount of purified cJun in buffer A, then sequentially dialyzed in buffer B (20 mM Tris (pH 7.9), 0.1 mM EDTA, 10% glycerol, 5 mM DTT) containing: (1) 1 M urea and 1 M NaCl, (2) 1 M NaCl, then (3) 0.1 M NaCl. Purified and folded cJun/cFos homodimers were aliquoted and stored at -80 °C.

Electrophoretic mobility shift assays

Recombinantly expressed cJun, cJun-bZip, or cFos was diluted into buffer DB (100 mM KCl, 20% glycerol, 40 mM Tris (pH 7.9), 0.1 mM DTT, 120 μg/ml BSA, and 0.12% NP-40). ^{32}P-Labeled aptamers were diluted into buffer RM (100 mM KCl, 8 mM $MgCl_2$, 40 mM HEPES (pH 7.9), 0.1 mM DTT), then heated to 95 °C for 2 minutes before being rapidly cooled in an ice bath to favor the most thermodynamically stable conformations. Reactions were assembled in 20 μL of buffer DB/RM (100 mM KCl, 10% glycerol, 4 mM MgCl2, 20 mM Tris (pH 7.9), 20 mM HEPES (pH 7.9), 0.1 mM DTT, 60 μg/ml BSA, and 0.06% NP-40) on ice, then the complexes were allowed to form by incubation for 20 minutes at room temperature. After incubation, 40 ng of poly dI•dC were added to each reaction as a non-specific competitor, and the reactions were subjected to electrophoresis through a 4% polyacrylamide gel containing 1x TG buffer and 5% glycerol. Gels were run for 60-90 minutes, then transferred to Whatman paper and dried. Radioactivity was detected using a phosphorimager screen (Amerisham Biosciences) then scanned using a Typhoon scanner. In all cases, DNA concentrations were kept at 50 pM.

S1 Nuclease footprinting

10 fmol of ^{32}P-labeled aptamer-19 were folded in buffer RM (100 mM KCl, 8 mM $MgCl_2$, 40 mM HEPES (pH 7.9), 0.1 mM DTT) by heating to 95 °C for 2 minutes before rapidly cooling in an ice bath. This was then added to 1x S1 Nuclease buffer (Promega), and incubated on ice with between 1 and 20 units of S1 Nuclease for 1 minute. The reaction was quenched using stop mix (200 mM KCl, 50 mM EDTA, 0.3 mg/mL yRNA, 1% glycogen), then phenol-chloroform extracted. The resulting fragments were precipitated in ethanol before resolving on an 11% denaturing polyacrylamide gel. Radioactivity was detected using a phosphorimager screen (Amerisham Biosciences) then scanned using a Typhoon scanner, and the data were quantitated using ImageJ software.

DNase1 footprinting

10 fmol of ^{32}P-labeled aptamer-19 were folded in buffer RM (100 mM KCl, 8 mM $MgCl_2$, 40 mM HEPES (pH 7.9), 0.1 mM DTT) by heating to 95 °C for 2 minutes before rapidly cooling in an ice bath. This was added to recombinant purified cJun or cJun-bZip in buffer DB/RM (100 mM KCl, 10% glycerol, 4 mM $MgCl_2$, 20 mM Tris (pH 7.9), 20 mM HEPES (pH 7.9), 0.1 mM DTT, 60 μg/ml BSA, and 0.06% NP-40) in a 20 μL reaction, and the complexes were allowed to form for 20 minutes at 30 °C. After 20 minutes, 5 units DNase and 20 μmol $CaCl_2$ were added to the reaction and incubated for 1 minute. The reaction was quenched using stop mix (200 mM KCl, 50 mM EDTA, 0.3 mg/mL yRNA, 1% glycogen), then phenol-chloroform extracted. The resulting fragments were precipitated in ethanol before resolving on an 11% denaturing polyacrylamide gel. Radioactivity was detected using a phosphorimager screen (Amerisham Biosciences) then scanned using a Typhoon scanner, and the data were quantitated using ImageJ software.

References

1. Carninci, P. et al. The transcriptional landscape of the mammalian genome. *Science* **309**, 1559–63 (2005).
2. Thomas, M. C., & Chiang, C. M. The general transcription machinery and general cofactors, *Crit. Rev. Biochem. Mol. Biol.* **41**, 105-178 (2006).
3. Cox, M. M, Doudna, J. A., O'Donnel, M. eds. Molecular Biology: Principles and Practice. W.H. Freeman and Company. New York (2012).
4. Sydow, J. F. & Cramer, P. RNA polymerase fidelity and transcriptional proofreading. *Curr. Opin. Struct. Biol.* **19**, 732–9 (2009).
5. Cramer, P. Architecture of RNA Polymerase II and Implications for the Transcription Mechanism. *Science*. **288**, 640–649 (2000).
6. Young, R. RNA polymerase II, *Annu. Rev. Biochem.* **60**, 689-715 (1991).
7. Buratowski, S. Progression through the RNA polymerase II CTD cycle, *Mol. Cell* **36**, 541-546 (2009).
8. Lee, T. I., & Young, R. A. Transcription of eukaryotic protein-coding genes, *Annu. Rev. Genet.* **34**, 77-137 (2000).
9. Matsui, T., Segall, J., Weil, P. A., & Roeder, R. G. Multiple factors required for accurate initiation of transcription by purified RNA polymerase II, *J. Biol. Chem.* **255**, 11992-11996 (1980).
10. Verrijzer, C. P., Chen, J.-L., Yokomori, K., & Tjian, R. Binding of TAFs to core elements directs promoter selectivity by RNA polymerase II, *Cell* **81**, 1115-1125 (1995).
11. Blair, R.H. Goodrich, J.A., & Kugel, J.F. Single molecule FRET shows uniformity in TBP-induced DNA bending and heterogeneity in bending kinetics. *Biochemistry* **51,** 7444-7455 (2012).
12. Smale, S. T., & Kadonaga, J. T. The RNA polymerase II core promoter, *Annu. Rev. Biochem.* **72**, 449-479 (2003).
13. Ohkuma, Y., & Roeder, R. G. Regulation of TFIIH ATPase and kinase activities by TFIIE during active initiation complex formation, *Nature* **368**, 160-163 (1994).
14. Holstege, F. C. P., van der Vliet, P. C., & Timmers, H. T. M. Opening of an RNA polymerase II promoter occurs in two distinct steps and requires the basal transcription factors IIE and IIH, *EMBO J.* **15**, 1666-1677 (1996).
15. Dvir, A., Conaway, J. W. & Conaway, R. C. Mechanism of transcription initiation and promoter escape by RNA polymerase II. *Curr. Opin. Genet. Dev.* **11**, 209–214 (2001).
16. Maston, G. A., Evans, S. K., & Green, M. R. Transcriptional regulatory elements in the human genome, *Annu. Rev. Genomics Hum. Genet.* **7**, 29-59 (2006).
17. Kadonaga, J. T. Regulation of RNA polymerase II transcription by sequence-specific DNA binding factors, *Cell* **116**, 247-257 (2004).
18. Spitz, F., & Furlong, E. E. M. Transcription factors: from enhancer binding to developmental control, *Nature Rev. Genet.* **13***,* 613-626 (2012)*.*
19. Knuesel, M. T., Meyer, K. D., Bernecky, C., & Taatjes, D. J. The human CDK8 subcomplex is a molecular switch that controls Mediator coactivator function, *Genes Dev.* **23**, 439-451 (2009).

20. McKenna, N. J. & O'Malley, B. W. Combinatorial control of gene expression by nuclear receptors and coregulators. *Cell* **108**, 465–74 (2002).
21. Diamond, M. I., Miner, J. N., Yoshinaga, S. K. & Yamamoto, K. R. Transcription factor interactions: selectors of positive or negative regulation from a single DNA element. *Science* **249**, 1266–72 (1990).
22. Grewal, S. I., & Moazed, D. Heterochromatin and epigenetic control of gene expression, *Science* **301**, 798-802 (2003).
23. Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F., & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 A resolution, *Nature* **389**, 251-260 (1997).
24. Kouzarides, T. Chromatin modifications and their function, *Cell* **128**, 693-705 (2007).
25. Campos, E. I., & Reinberg, D. Histones: annotating chromatin, *Annu. Rev. Genet.* **43**, 559-599 (2009).
26. Clapier, C. R., & Cairns, B. R. The biology of chromatin remodeling complexes, *Annu. Rev. Biochem.* **78**, 273-304 (2009).
27. Glover, J. N., & Harrison, S. C. Crystal structure of the heterodimeric bZIP transcription factor cFos-cJun bound to DNA, *Nature* **373**, 257-261 (1995).
28. Shaulian, E., & Karin, M. AP-1 as a regulator of cell life and death, *Nature Cell Bio.* **4**, E131-E136 (2002).
29. Eferl, R., & Wagner, E. F. AP-1: A Double-Edged Sword in Tumorigenesis. *Nat. Rev. Cancer* **3**, 859–68 (2003).
30. Karin, M. The regulation of AP-1 activity by mitogen-activated protein kinases, *J. Biol. Chem.* **270**, 16483-16486 (1995).
31. Karin, M., Liu, Z., & Zandi, E. AP-1 function and regulation, *Curr. Opin. Cell Biol.* **9**, 240-246 (1997).
32. Shaw, J. P., Utz, P. J., Durand, D. B., Toole, J. J., Emmel, E. A., & Crabtree, G. R. Identification of a putative regulator of early T cell activation genes, *Science* **241**, 202-205 (1988).
33. Hogan, P. G., Chen, L., Nardone, J. & Rao, A. Transcriptional regulation by calcium, calcineurin, and NFAT. *Genes Dev*. **17**, 2205–32 (2003).
34. Müller, M. R. & Rao, A. NFAT, immunity and cancer: a transcription factor comes of age. *Nat. Rev. Immunol.* **10**, 645–56 (2010).
35. Mancini, M. & Toker, A. NFAT proteins: emerging roles in cancer progression. *Nat. Rev. Cancer* **9**, 810–20 (2009).
36. Jauliac S, Lopez-Rodriguez C, Shaw LM, Brown LF, Rao A, & Toker A. The role of NFAT transcription factors in integrin-mediated carcinoma invasion. *Nat Cell Biol* **4**: 540–544 (2002).
37. Yiu, G. K. & Toker, A. NFAT induces breast cancer cell invasion by promoting the induction of cyclooxygenase-2. *J. Biol. Chem.* **281**, 12210–7 (2006).
38. Macian, F. NFAT proteins: key regulators of T-cell development and function, *Nat. Rev. Immunol.* **5**, 472-484 (2005).
39. Macián, F., López-Rodríguez, C. & Rao, a. Partners in transcription: NFAT and AP-1. *Oncogene* **20**, 2476–89 (2001).
40. Ali, S. & Coombes, R.C. Endocrine-responsive breast cancer and strategies for combating resistance. *National Review of Cancer* **2**, 101-112 (2002).
41. Hanahan, D. & Weinberg, R. a. Hallmarks of cancer: the next generation. *Cell* **144**, 646–74 (2011).

42. Tannheimer, S. L., Ethier, S. P., Caldwell, K. K. & Burchiel, S. W. Benzo[a]pyrene- and TCDD-induced alterations in tyrosine phosphorylation and insulin-like growth factor signaling pathways in the MCF-10A human mammary epithelial cell line. *Carcinogenesis* **19**, 1291–7 (1998).
43. Tannheimer, S. L., Barton, S. L., Ethier, S. P. & Burchiel, S. W. Carcinogenic polycyclic aromatic hydrocarbons increase intracellular Ca2+ and cell proliferation in primary human mammary epithelial cells. *Carcinogenesis* **18**, 1177–82 (1997).
44. Siewit, C. L., Gengler, B., Vegas, E., Puckett, R. & Louie, M. C. Cadmium promotes breast cancer cell proliferation by potentiating the interaction between ERalpha and c-Jun. *Mol. Endocrinol*. **24**, 981–92 (2010).
45. Shimada, T. & Fujii-Kuriyama, Y. Metabolic activation of polycyclic aromatic hydrocarbons to carcinogens by cytochromes P450 1A1 and 1B1. *Cancer Sci.* **95**, 1–6 (2004).
46. Wang, S. H., Shih, Y. L., Ko, W. C., Wei, Y. H. & Shih, C. M. Cadmium-induced autophagy and apoptosis are mediated by a calcium signaling pathway. *Cell. Mol. Life Sci.* **65**, 3640–52 (2008).
47. Byrne, C., Divekar, S. D., Storchan, G. B., Parodi, D. a & Martin, M. B. Metals and breast cancer. *J. Mammary Gland Biol. Neoplasia* **18**, 63–73 (2013).
48. Tannheimer, S. L., Lauer, F. T., Lane, J. & Burchiel, S. W. Factors influencing elevation of intracellular Ca2+ in the MCF-10A human mammary epithelial cell line by carcinogenic polycyclic aromatic hydrocarbons. *Mol. Carcinog.* **25**, 48–54 (1999).
49. Waisberg, M., Joseph, P., Hale, B. & Beyersmann, D. Molecular and cellular mechanisms of cadmium carcinogenesis. *Toxicology* **192**, 95–117 (2003).
50. Palmer, A. E. et al. Ca2+ indicators based on computationally redesigned calmodulin-peptide pairs. *Chem. Biol.* **13**, 521–30 (2006).
51. McCombs, J. E. & Palmer, A. E. Measuring calcium dynamics in living cells with genetically encodable calcium indicators. *Methods* **46**, 152–9 (2008).
52. Chatila, T., Silverman, L., Miller, R. & Geha, R. Mechanisms of T cell activation by the calcium ionophore ionomycin. *J. Immunol.* **143**, 1283–9 (1989).
53. Shaner, N. C. et al. Improving the photostability of bright monomeric orange and red fluorescent proteins. *Nat. Methods* **5**, 545–551 (2008).
54. Walters, R. D., Drullinger, L. F., Kugel, J. F. & Goodrich, J. A. NFATc2 recruits cJun homodimers to an NFAT site to synergistically activate interleukin-2 transcription. *Mol. Immunol.* **56**, 48–56 (2013).
55. Iñiguez, M. a, Martinez-Martinez, S., Punzón, C., Redondo, J. M. & Fresno, M. An essential role of the nuclear factor of activated T cells in the regulation of the expression of the cyclooxygenase-2 gene in human T lymphocytes. *J. Biol. Chem.* **275**, 23627–35 (2000).
56. 1. Chen, M. & O'Connor, K. L. Integrin alpha6beta4 promotes expression of autotaxin/ENPP2 autocrine motility factor in breast carcinoma cells. Oncogene 24, 5125–30 (2005).
57. Larkins, T. L., Nowell, M., Singh, S. & Sanford, G. L. Inhibition of cyclooxygenase-2 decreases breast cancer cell motility, invasion and matrix metalloproteinase expression. *BMC Cancer* **6**, 181 (2006).
58. Rao, a, Luo, C. & Hogan, P. G. Transcription factors of the NFAT family: regulation and function. *Annu. Rev. Immunol.* **15**, 707–47 (1997).

59. Dolmetsch, R. E., Xu, K. & Lewis, R. S. Calcium oscillations increase the efficiency and specificity of gene expression. *Nature* **392**, 933–6 (1998).
60. Nguyen, T. N., Kim, L. J., Walters, R. D., Drullinger, L. F., Lively, T. N., Kugel, J. F., & Goodrich, J. A. The C-terminal region of human NFATc2 binds cJun to synergistically activate interleukin-2 transcription, *Mol*. *Immunol*. **47**, 2314-2322 (2010).
61. Rhee, H.-W. et al. Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. *Science* **339**, 1328–31 (2013).
62. Martell, J. D. et al. Engineered ascorbate peroxidase as a genetically encoded reporter for electron microscopy. *Nat. Biotechnol.* **30**, 1143–8 (2012).
63. Krivickas, S. J., Tamanini, E., Todd, M. H. & Watkinson, M. Effective methods for the biotinylation of azamacrocycles. *J. Org. Chem.* **72**, 8280–9 (2007).
64. Stoltenburg, R., Reinemann, C. & Strehlitz, B. SELEX--a (r)evolutionary method to generate high-affinity nucleic acid ligands. *Biomol. Eng.* **24**, 381–403 (2007).
65. Walters, R.D. Mechanisms of transcriptional regulation by NFAT and AP-1 transcriptional activators. *University of Colorado Boulder, Department of Chemistry and Biochemistry*. Doctoral thesis (2013).
66. Carrillo, R. J., Dragan, a I. & Privalov, P. L. Stability and DNA-binding ability of the bZIP dimers formed by the ATF-2 and c-Jun transcription factors. *J. Mol. Biol.* **396**, 431–40 (2010).
67. Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**, 3406–3415 (2003).
68. SantaLucia, Jr. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl. Acad. Sci. USA* **95**, 1460-1465 (1998).